Charles University in Prague
Faculty of Social Sciences
Institute of Economic Studies

Three Essays on Economic Dynamics

Dissertation Thesis

2009 Petr Kadeřábek

Title:	Three Essays on Economic Dynamics
Author:	Petr Kadeřábek
Advisor:	prof. Jan Kodera
Affiliation:	Charles University in Prague, Faculty of Social Sciences, Institute of Economic Studies
Year:	2009

I hereby declare that this dissertation is my original work and that all sources are properly identified and acknowledged.

Plzeň, January 30, 2009

Acknowledgements

I wish to thank my advisor prof. Jan Kodera for his good advices and willingness to swiftly discuss my papers, and to Josef Vodička, Aleš Slabý, prof. Jiří Reif, the anonymous referees, and participants of the UTIA seminar for valuable suggestions. Many thanks go to Míša and my family for patience with my studies. Financial support for this research from Czech Science Foundation, project Nonlinear Dynamic Economic Systems under No. GAČR 402/03/H057 is gratefully acknowledged.

Contents

Foreword

The dissertation consists of three essays on dynamic models in economics. First two of them present particular models, the third one proposes a method how to use the dynamic models effectively for prediction of the systems, which possess unpredictable dynamics in the long term.

In Essay 1 *Price Gap: Complement of the Inventory Gap*, we show that both the excess supply and the excess supply price are given by two variables: the inventory gap and the price gap defined as the gap between the sales price and the sellers' costs incl. the normal profit. These measures of disequilibrium between supply and demand are used in the Walrasian and Marshallian adjustment processes. The main contribution of this paper is the specification of the price gap, which can be also successfully used in the empirical models as shown in the application on the panel data of European Union countries.

This essay is based on the paper, published as P. Kadeřábek, Simple model of interaction between CPI and PPI, *Czech Journal of Political Economy*, 2, 2006.

Essay 2 *Stress Testing Probability of Default of Individuals* introduces a model for stress testing of probability of default of individuals. The model rests on assumption that the individual defaults if his income plus savings do not cover his consumption and installments. The probability of default is then described as a function of several macroeconomic indicators: interest rates, wages, unemployment, and price level. Stress testing is carried out by applying exogenous stress scenarios for development of these indicators. The model implies that sensitivity of probability of default to the stress is mainly driven by Installment to Income Ratio and for mortgages also by loan maturity: These tools are suggested as appropriate to manage credit risk of retail portfolios.

This essay is the modified version of the paper P. Kadeřábek, A. Slabý,

J. Vodička, Stress Testing of Probability of Default of Individuals, published in *Prague Economic Papers* 4/2008.

In Essay 3 *Correcting Predictive Models of Chaotic Systems*, we present a method for combination of two predictive models of a chaotic system: A short-term model, usually performing iterated one-step predictions, is combined (corrected) with a natural measure of the system. The resulting model is given uniquely and, for any prediction horizon, it dominates both the short-term model and the natural measure (in the sense of mean square error). To allow for a simple application, the Monte Carlo approach for approximating distribution of prediction of the short-term model is discussed, as well as a nonparametric specification of the natural measure. Closed-form formula of the resulting model is derived for AR(1) short-term model and a Gaussian natural measure. Effectiveness of the correction method is demonstrated on the U.S. unemployment rate and the NMR laser data set.

Early version of this essay was published in the IES Working Papers as P. Kadeřábek, Correcting Predictive Models of Chaotic Reality, *IES Working Papers*, 31, 2006.

Essay 1

Price Gap: Complement of the Inventory Gap

1.1 Introduction

We assume that we can observe one price-quantity pair on the sellers' supply curve and one on the buyers' demand curve. The usual theoretical measures of disequilibrium are the excess supply and the excess supply price. We show that both of these measures are fully determined by the inventory gap and the price gap. The inventory gap, defined as the gap between the actual and desired inventories, is the traditional topic in the literature, see e.g. Blinder and Maccini (1991). The main contribution of this paper is the price gap, defined as the gap between the sales price and the sellers' costs (incl. the normal profit).

The econometric application demonstrated in this paper is the following: We consider that the sellers are traders, who buy the production at the production costs (measured by the producer price index) and distribute it. Then, the sellers' price consists of the producer price (observable) and the distribution costs (is assumed to change relatively slowly hence filtered out), the quantity is measured by the volume index of production. The price-quantity pair of the buyers is the consumer price index and volume of retail sales.

The excess supply derived from the inventory gap and price gap is used in the dynamic model based on the Walrasian adjustment mechanism. Similarly, the excess supply price is employed in the dynamic model of the Marshallian type. An equivalence of both models is shown under the assumption of a positively-sloped supply curve and a negatively-sloped demand curve. Note that equivalence of the models based on Walrasian and Marshallian adjustment mechanisms does not naturally mean equivalence of these adjustment mechanisms.

The dynamic models are empirically verified on the panel data of European Union countries.

1.2 Theoretical framework

We will follow the notation of Blinder and Maccini (1991).

The supply of the sellers is of the form

$$Y_t - N_t^* + N_{t-1} = S_t(P_t^y), \tag{1.1}$$

where Y_t is real output for manufacturers or purchases for wholesalers and retailers, N_{t-1} the stock of inventories inherited from the previous period, and N_t^* denotes the target level of inventories. The price P_t^y is determined by the sellers' costs including the normal profit. This point on the supply curve is observed or at least can be estimated.

The demand of the buyers

$$X_t = D_t(P_t^x) \tag{1.2}$$

returns the quantity X_t bought for the price P_t^x. This is the only directly observable point on the demand curve.

Consider the identity, used e.g. by Blinder and Maccini (1991),

$$N_t - N_{t-1} \equiv Y_t - X_t, \tag{1.3}$$

hence the inventory gap can be derived as

$$S_t(P_t^y) - D_t(P_t^x) = Y_t - N_t^* + N_{t-1} - X_t = N_t - N_t^*. \tag{1.4}$$

However this difference between supply and demand concerns to different prices. When using the same price, expression for the excess supply [segments (A-A') and (B-B') from Fig. 1.1] can be written as

$$S_t(P_t^x) - D_t(P_t^x) = [S_t(P_t^y) - D_t(P_t^x)] + [S_t(P_t^x) - S_t(P_t^y)], \tag{1.5}$$

$$S_t(P_t^y) - D_t(P_t^y) = [S_t(P_t^y) - D_t(P_t^x)] + [D_t(P_t^x) - D_t(P_t^y)]. \tag{1.6}$$

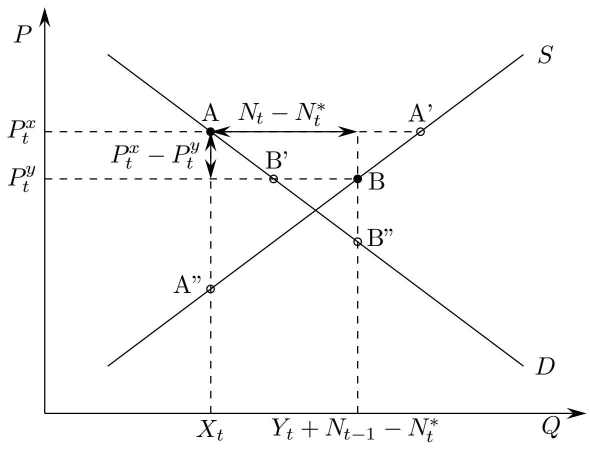

Figure 1.1: Measures of disequilibrium.

To be able to simplify expressions of gaps, assume linear S_t and D_t,

$$S_t(P) = s_{0,t} + s_1 P, \tag{1.7}$$
$$D_t(P) = d_{0,t} - d_1 P, \tag{1.8}$$

where $s_1, d_1 > 0$. The curves may shift due to the dependence of $s_{0,t}$ and $d_{0,t}$ on time, while their slopes given by s_1 and d_1 are constant over time. The term in the first square brackets of Eq. (1.5) and (1.6) is equal to $N_t - N_t^*$ by Eq. (1.4). The terms in the second square brackets can be simplified to the multiples of $P_t^x - P_t^y$ due to Eq. (1.7) and (1.8). Then the excess supply is

$$S_t(P_t^x) - D_t(P_t^x) = (N_t - N_t^*) + s_1(P_t^x - P_t^y), \tag{1.9}$$
$$S_t(P_t^y) - D_t(P_t^y) = (N_t - N_t^*) - d_1(P_t^x - P_t^y). \tag{1.10}$$

We also need the excess supply prices. It is easy to state analogy of Eq. (1.4),

$$S_t^{-1}(Y_t - N_t^* + N_{t-1}) - D_t^{-1}(X_t) = -(P_t^x - P_t^y), \tag{1.11}$$

and derive the price differences for the same quantities, which refer to lengths of the segments (A-A") and (B-B"),

$$S_t^{-1}(X_t) - D_t^{-1}(X_t) = -s_1^{-1}(N_t - N_t^*) - (P_t^x - P_t^y), \tag{1.12}$$

$$S_t^{-1}(Y_t - N_t^* + N_{t-1}) - D_t^{-1}(Y_t - N_t^* + N_{t-1}) \\ = d_1^{-1}(N_t - N_t^*) - (P_t^x - P_t^y). \tag{1.13}$$

Note that P_t^x and P_t^y can equivalently be defined as *log* prices without any impact on the meaning of the equations. However, this is not the case of X_t, Y_t, N_t^*, or N_{t-1}: Identity (1.3) does not hold for *log* quantities.

1.3 Equilibrium adjustment

We present two models in this section: The first one, based on the Walrasian adjustment mechanism, which uses the excess supply quantities given by Eq. (1.9) and (1.10), and the other one, employing the excess supply prices from Eq. (1.12) and (1.13) in the Marshallian adjustment process. In both models, the sellers and buyers adjust separately to the equilibrium.

The principle of the Walrasian adjustment is that a positive excess supply always leads to a decline in price and vice versa. In the Marshallian adjustment, a positive excess supply price causes a decline in the quantity and vice versa. For the positively-sloped supply curve and negatively-sloped demand curve ($s_1, d_1 > 0$ in our model), both approaches lead to stable equilibrium. However, in other cases, stability of the equilibrium depends on the adjustment mechanism.

There is not consesus in the literature, where the Marshallian or Walrasian approach suits better. Takayama (1985), p. 299, states that "the Marshallian stability conditions are explicitly designed for the theory of production, whereas the Walrasian price adjustment is more suited for the theory of exchange." Blaug (1997), p. 391, disagrees with such statements and argues that "in the short run, when stocks of goods are not given and output can be varied from given plants and equipment, the output-adjuster model is just as reasonable as the price-adjuster model."

We find the one-to-one mapping between the parameters of both presented models for the typical case of the positively-sloped supply curve and negatively-sloped demand curve, hence they show to be equivalent and there is no need to discuss which alternative to choose.

1.3.1 Walrasian adjustment

The Walrasian adjustment mechanism in the discrete time is

$$P_{t+1} - P_t = -\alpha_w \left[S(P_t) - D(P_t)\right], \tag{1.14}$$

where $\alpha_w > 0$ is the speed of adjustment. Price P_t adjusts according to the difference between the quantity supplied $S(P_t)$ and quantity demanded $D(P_t)$. Note that *no uniquely given quantity exists.*

The Walrasian adjustment of P_t^x and P_t^y, augmented for price expectations P_t^{x*} (price expected by the buyers) and P_t^{y*} (price expected by the sellers), is described by

$$P_{t+1}^x - P_{t+1}^{x*} = -\alpha_{w,x} \left[S_t(P_t^x) - D_t(P_t^x)\right], \tag{1.15}$$

$$P_{t+1}^y - P_{t+1}^{y*} = -\alpha_{w,y} \left[S_t(P_t^y) - D_t(P_t^y)\right], \tag{1.16}$$

$\alpha_{w,x}, \alpha_{w,y} > 0$ denote speeds of adjustment. The supply and demand is assumed to depend on the plans Y_t^* (real output planned by the sellers) and X_t^* (quantity planned by the buyers to be bought), and the difference between the actual and expected price level, hence the precised version of Eq. (1.7) and (1.8) is

$$D_t(P_t^x) = X_t^* - d_1(P_t^x - P_t^{x*}), \tag{1.17}$$

$$S_t(P_t^y) = Y_t^* - N_t^* + N_{t-1} + s_1(P_t^y - P_t^{y*}). \tag{1.18}$$

From the pairs of Eq. (1.1) and (1.18) and similarly Eq. (1.2) and (1.17), we obtain formulas for the deviation of X_{t+1} and Y_{t+1} from the plans,

$$X_{t+1} - X_{t+1}^* = -d_1(P_{t+1}^x - P_{t+1}^{x*}), \tag{1.19}$$

$$Y_{t+1} - Y_{t+1}^* = s_1(P_{t+1}^y - P_{t+1}^{y*}). \tag{1.20}$$

Equation (1.20) is in fact the Lucas supply function, see Lucas (1976). It was extended in Blinder and Fischer (1982) by adding the difference between the actual and desired inventories. We use another way to introduce inventories to the equations: We perform substitution from Eq. (1.15) and (1.16) and obtain

$$X_{t+1} - X_{t+1}^* = d_1\alpha_{w,x} \left[S_t(P_t^x) - D_t(P_t^x)\right], \tag{1.21}$$

$$Y_{t+1} - Y_{t+1}^* = -s_1\alpha_{w,y} \left[S_t(P_t^y) - D_t(P_t^y)\right]. \tag{1.22}$$

In Sec. 1.3.3, the inventory and price gaps will be substituted to the right hand side of the equations.

1.3.2 Marshallian adjustment

The Marshallian adjustment mechanism is described by the equation

$$Q_{t+1} - Q_t = -\alpha_m \left[S^{-1}(Q_t) - D^{-1}(Q_t) \right], \tag{1.23}$$

where $\alpha_m > 0$ is the speed of adjustment. Quantity Q_t adjusts according to the difference between the supply price and demand price. In this case, *no uniquely given price exists*.

Analogically to the previous section, dynamics of the Marshallian adjustment of X_t and Y_t augmented for the plans, is described by

$$X_{t+1} - X^*_{t+1} = -\alpha_{m,x} \left[S_t^{-1}(X_t) - D_t^{-1}(X_t) \right], \tag{1.24}$$

$$Y_{t+1} - Y^*_{t+1} = -\alpha_{m,y} \left[S_t^{-1}(Y_t - N_t^* + N_{t-1}) - D_t^{-1}(Y_t - N_t^* + N_{t-1}) \right], \tag{1.25}$$

$\alpha_{m,x}, \alpha_{m,y} > 0$ denote speeds of adjustment. Inverting Eq. (1.19) and (1.20) yields

$$P^x_{t+1} - P^{x*}_{t+1} = -d_1^{-1}(X_{t+1} - X^*_{t+1}), \tag{1.26}$$

$$P^y_{t+1} - P^{y*}_{t+1} = s_1^{-1}(Y_{t+1} - Y^*_{t+1}). \tag{1.27}$$

After substitution of Eq. (1.24) and (1.25) into (1.26) and (1.27),

$$P^x_{t+1} - P^{x*}_{t+1} = d_1^{-1}\alpha_{m,x} \left[S_t^{-1}(X_t) - D_t^{-1}(X_t) \right], \tag{1.28}$$

$$P^y_{t+1} - P^{y*}_{t+1} = -s_1^{-1}\alpha_{m,y} \left[S_t^{-1}(Y_t + N_{t-1} - N_t^*) - D_t^{-1}(Y_t + N_{t-1} - N_t^*) \right]. \tag{1.29}$$

1.3.3 Common inference

Both ajdustment mechanisms, described by Eq. (1.14) and (1.23), have *different state variables* (without a unique relationship between them), hence they *cannot be directly compared.* However, the dynamic models based on them have the same state variables, hence can be compared and we show equivalence of both approaches under the typically-sloped supply and demand curves.

After substitution from Eq. (1.9) and (1.10), equations from Sec. 1.3.1

for the Walrasian adjustment can be written as

$$
\begin{aligned}
P^x_{t+1} - P^{x*}_{t+1} &= -\alpha_{w,x}(N_t - N^*_t) - \alpha_{w,x}s_1(P^x_t - P^y_t), &(1.30)\\
P^y_{t+1} - P^{y*}_{t+1} &= -\alpha_{w,y}(N_t - N^*_t) + \alpha_{w,y}d_1(P^x_t - P^y_t), &(1.31)\\
X_{t+1} - X^*_{t+1} &= \alpha_{w,x}d_1(N_t - N^*_t) + \alpha_{w,x}s_1d_1(P^x_t - P^y_t), &(1.32)\\
Y_{t+1} - Y^*_{t+1} &= -\alpha_{w,y}s_1(N_t - N^*_t) + \alpha_{w,y}s_1d_1(P^x_t - P^y_t). &(1.33)
\end{aligned}
$$

Analogically, equations of the Marshallian adjustment from Sec. 1.3.2 become after substitution from Eq. (1.12) and (1.13)

$$
\begin{aligned}
P^x_{t+1} - P^{x*}_{t+1} &= -\alpha_{m,x}s_1^{-1}d_1^{-1}(N_t - N^*_t) - \alpha_{m,x}d_1^{-1}(P^x_t - P^y_t), &(1.34)\\
P^y_{t+1} - P^{y*}_{t+1} &= -\alpha_{m,y}s_1^{-1}d_1^{-1}(N_t - N^*_t) + \alpha_{m,y}s_1^{-1}(P^x_t - P^y_t), &(1.35)\\
X_{t+1} - X^*_{t+1} &= \alpha_{m,x}s_1^{-1}(N_t - N^*_t) + \alpha_{m,x}(P^x_t - P^y_t), &(1.36)\\
Y_{t+1} - Y^*_{t+1} &= -\alpha_{m,y}d_1^{-1}(N_t - N^*_t) + \alpha_{m,y}(P^x_t - P^y_t). &(1.37)
\end{aligned}
$$

We can see that the Walrasian and Marshallian equations are equivalent, when $\alpha_{m,y} = \alpha_{w,y}s_1d_1$ and $\alpha_{m,x} = \alpha_{w,x}s_1d_1$. This holds for the case of the positively-sloped supply curve and negatively-sloped demand curve, i.e. $s_1, d_1 > 0$. If one of the curves has an opposite slope then either $\alpha_{w,y} < 0$ or $\alpha_{m,y} < 0$ (resp. either $\alpha_{w,x} < 0$ or $\alpha_{m,x} < 0$), which violates the Walrasian or Marshallian adjustment principle.

1.4 Empirical implementation

We implement the model empirically in this section. Lot has been written about the inventory gap, hence we focus on the verification of what is new in this paper: the price gap defined as a gap between different price indices.

1.4.1 Data

Consider the sellers to be the traders, who buy the produced consumer goods at the production costs and distribute them to the consumers at the consumer prices. The most important problem is in specification of P^y_t, which consists of the producer prices and the unobservable distribution costs incl. the normal profit. The distribution costs are assumed to change relatively slowly, hence may be estimated using the filter, which also adjusts slowly to the observed time series of the difference between the consumer and producer prices.

All the data used are provided by Eurostat:

P^x	log of harmonized consumer price index (non-energy industrial goods)
P^y	log of producer price index (manufacturing)
X	log of retail trade deflated turnover index (household goods, data adjusted by working days)
Y	log of production index (manufacturing, data adjusted by working days)

The data concern to the panel of EU countries as of May, 1, 2004, when the eight countries of central and eastern Europe joined EU. Estonia, Latvia, and Malta were removed, because P^y is not available for them. Luxembourg was omitted due to its substantially higher volatility of quarterly producer price inflation ΔP^y (standard deviation of 0.0394 compared to 0.0124 of the rest of the panel).

Most time series come from the period from Q1 1990 to Q2 2008. However, their usability in regression is limited by the P^x data, which are available since Q1 1996.

Note that we assume P^x, P^y, X, and Y to be in logs throughout Sec. 1.4: We have already mentioned in Sec. 1.2 that P_t^x and P_t^y can equivalently be defined as *log* prices. Sensitivity of X, resp. Y, to the price gap is assumed to be proportional to X^*, resp. Y^*, which is treated by taking the logs of X and Y.

1.4.2 Price gap estimation

We assume in our model that all the variables X_t, Y_t, P_t^x, and P_t^y concern to the same group of "consumer goods". However, this is not precisely true for the empirical variables. Goods that are present in only one of the price indices may be subject to different inflation, hence cause nonstationarity of $P_t^x - P_t^y$.

To verify plausibility of using directly $P_t^x - P_t^y$ as a price gap, we test its stationarity with the tests by Hylleberg et al. (1990) (HEGY test), which is designed for the seasonal data, and Kwiatkowski et al. (1992) (KPSS test), having a null hypothesis of stationarity.

We choose 4 lags in the KPSS test in line with Ghysels and Perron (1993), who suggest to use the number of lags greater or equal to the number of seasons when working with the seasonal data.

According to Ghysels et al. (1994), we introduce the dummy variables for the seasons when using the HEGY test. We use 1 lag only, as the test is designed for the seasonal data, hence we do not have to deal with them using more lags. Moreover, lower number of lags increases power of the test and thus possible non-rejection of the null hypothesis will be a stronger evidence of non-stationarity.

In both tests, data are assumed to have a constant deterministic component. The KPSS test rejects the null hypothesis of stationarity on the 5% level for 20 out of 21 countries. The HEGY test does not reject the null hypothesis of non-stationarity in the case of any of 21 countries on the 10% level. These results show strong non-stationarity, hence filtering of $P_t^x - P_t^y$ is necessary.

Assume that the non-stationarity of $P_t^x - P_t^y$ is caused by the slowly-changing trend. To estimate it, we use the filter described in Hodrick and Prescott (1997), refered to as the HP filter.

The HP filter is constructed as two-sided, i.e. both past and future observations are used to compute the trend value in some point. This property could introduce the spurious regression to the equations explaining the CPI and PPI inflation. Therefore, we do not use future variables to compute the trend and use the HP filter as one-sided.

The HP filter has a parameter λ. For $\lambda = 0$, the trend is equal to the filtered time series, while for $\lambda = \infty$, the filter is equivalent to fitting the linear trend by the ordinary least sqares method. The value of $\lambda = 1600$ is suggested for the quarterly data and is used in this paper.

We have decided that the time series of at least 8 observations is required for filtering, hence the trend is not computed for the 7 values at the beginning of the time series.

The filtered price gap is seasonaly adjusted by the moving average computed from the last 4 values.

1.4.3 Regression equations

Denote π_t the price gap estimated in line with Sec. 1.4.2. Inflation and gaps of production and consumption are defined

$$p_t^x \equiv P_t^x - P_{t-1}^x, \tag{1.38}$$
$$p_t^y \equiv P_t^y - P_{t-1}^y, \tag{1.39}$$
$$x_t \equiv X_t - X_t^*, \tag{1.40}$$
$$y_t \equiv Y_t - Y_t^*, \tag{1.41}$$

where X_t^* and Y_t^* were obtained by the same filter as was the price gap (see Sec. 1.4.2) but without the seasonal adjustment. Seasonality is treated differently in this case.

We assume regressive inflation expectations: The expected inflation in time $t-1$ for time t is equal to the weighted average of inflation from the same season of the previous year and the equilibrium inflation $\overline{p}$,

$$P_t^* - P_{t-1} = (1-\rho)(P_{t-4} - P_{t-5}) + \rho\overline{p}, \tag{1.42}$$

where $\rho \in [0,1]$. The equilibrium inflation $\overline{p}$ can be interpreted as the inflation target of the central bank. For $\rho = 0$, the expectations become static. Hence, Eq. (1.42) implies $P_t - P_t^* = p_t - (1-\rho)p_{t-4} - \rho\overline{p}$. This specification has econometric advantages: It introduces the lagged inflation to the equation, hence captures seasonality in inflation and is a proxy for the predictive variables other than the price gap. From the same reason, we incorporate the lagged dependent variables also into the equations of production and sales. The regression equations are

$$p_t^x = \beta_{1,0} + \beta_{1,1}p_{t-4}^x + \beta_{1,2}\pi_{t-4} + \varepsilon_{1,t}, \tag{1.43}$$
$$p_t^y = \beta_{2,0} + \beta_{2,1}p_{t-4}^y + \beta_{2,2}\pi_{t-4} + \varepsilon_{2,t}, \tag{1.44}$$
$$x_t = \beta_{3,0} + \beta_{3,1}x_{t-4} + \beta_{3,2}\pi_{t-4} + \varepsilon_{3,t}, \tag{1.45}$$
$$y_t = \beta_{4,0} + \beta_{4,1}y_{t-4} + \beta_{4,2}\pi_{t-4} + \varepsilon_{4,t}, \tag{1.46}$$

where $\varepsilon_{1,t}$, $\varepsilon_{2,t}$, $\varepsilon_{3,t}$, and $\varepsilon_{4,t}$ are disturbances. We expect $\beta_{1,2} < 0$ and $\beta_{2,2}, \beta_{3,2}, \beta_{4,2} > 0$.

Regression equations are estimated by the fixed-effects model for panel data.

The Breusch-Godfrey test of autocorrelation of residuals up to lag 4 is performed. [For the test in univariate regression, see Breusch (1978) and

	Estimate	t-stat. (p-val.)	Rob. t (p-val.)	DF
$\beta_{1,2}$	0.1072	3.7022 (0.0001)	2.8424 (0.0023)	641
$\beta_{2,2}$	−0.0440	−3.0618 (0.0010)	−3.6921 (0.0001)	641
$\beta_{3,2}$	0.3107	3.8512 (<.0001)	2.4881 (0.0065)	628
$\beta_{4,2}$	0.4431	2.5700 (0.0050)	1.8924 (0.0294)	586

Table 1.1: Estimates of the coefficients of the price gap: p-values refer to the null hypothesis $\beta_{1,2} < 0$, resp. $\beta_{i,2} > 0$ for $i \in \{2, 3, 4\}$.

Godfrey (1978), its use for panel data is shown in Wooldridge (2002), p.288.] The test strongly rejects the null hypothesis of no autocorrelation, hence the t-statistic of the coefficients, based on the estimate of the standard error robust to heteroskedasticity and autocorrelation, is included.

Estimates of the coefficients of the price gap from Eq. (1.43)–(1.46) and their statistical significances are shown in Table 1.1: The classical t-test rejects the null hypothesis of zero or wrong sign of the coefficient in all 4 equations on the 1% level. These results are biased due to autocorrelation of residuals. According to the robust t-test, the price gap is significant on the 5% level in the equation of production and significant on the 1% level in the remaining 3 cases. These empirical findings support our hypotheses.

1.5 Conclusion

We have shown that the excess supply (at the same price) is given by both the inventory gap and the price gap, which was employed in the adjustment model of Walrasian type. The same factors determine the excess supply price (at the same quantity), used in the Marshallian adjustment model. Equivalence of both approaches under the condition of a positively-sloped supply curve and a negatively-sloped demand curve was shown.

The model was empirically verified on the panel of European Union countries. We have considered the market, where the sellers buy consumer goods for the producer prices and distribute them for the consumer prices. The amount supplied was given by production, the amount demanded by retail sales.

In the empirical implementation, we have focused on the verification of statistical significance of the price gap defined as a gap between different price indices, because it is the main contribution of this paper and there is already a large amount of literature about the inventory gap. The classical t-test rejected the null hypothesis of zero or wrong sign of the price gap coefficient in all 4 equations on the 1% level. However, these results are biased due to autocorrelation of residuals. According to the robust t-test, the price gap is significant on the 5% level in the equation of production and significant on the 1% level in the remaining 3 cases. These results support validity of the presented theoretical models.

Bibliography

M. Blaug. *Economic Theory in Retrospect*. Cambridge University Press, 1997.

A. Blinder and S. Fischer. Inventories, rational expectations, and the business cycle, 1982.

A. Blinder and L. Maccini. Taking stock: A critical assessment of recent research on inventories. *Journal of Economic Perspectives*, 5(1):73–96, 1991.

T. Breusch. Testing for autocorrelation in dynamic linear models. *Australian Economic Papers*, 17(31):334–355, 1978.

E. Ghysels and P. Perron. The effect of seasonal adjustment filters on tests for a unit root. *Journal of Econometrics*, 55(1-2):57–98, 1993.

E. Ghysels, H. Lee, and J. Noh. Testing for unit roots in seasonal time series. *Journal of Econometrics*, 62(2):415–442, 1994.

L. G. Godfrey. Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables. *Econometrica*, 46(6):1293–1301, November 1978.

R. Hodrick and E. Prescott. Postwar US business cycles: An empirical investigation. *Journal of Money, Credit & Banking*, 29(1):1–16, 1997.

S. Hylleberg, R. Engle, C. Granger, and B. Yoo. Seasonal integration and cointegration. *Journal of Econometrics*, 44(1-2):215–238, 1990.

D. Kwiatkowski, P. Phillips, P. Schmidt, and Y. Shin. Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54(1-3):159–178, 1992.

R. Lucas. Econometric policy evaluation: A critique. *Carnegie-Rochester Conference Series on Public Policy*, 1(2):19–46, 1976.

A. Takayama. *Mathematical Economics*. Cambridge University Press, 1985.

J. Wooldridge. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, 2002.

Essay 2

Stress Testing Probability of Default of Individuals

2.1 Introduction

Complex credit risk management in financial institutions addresses two levels of credit risk: expected losses and unexpected losses. Expected losses, that are long-term average losses, are covered through proper credit risk pricing and profitability management. Possible unexpected losses, that are higher than long-term average losses, are covered by capital reserves up to certain level of severity. Unexpected losses are rare by definition but the institution should foresee impact of extreme events, which could evolve from current situation, and have a plan for each such event to mitigate its impact if it evolves. This task is supported by stress testing.

Stress testing is, in other words, investigation of impact of meaningfully defined scenarios of future development, extreme development in particular. In credit risk stress testing, impact of these scenarios is measured in terms of losses to be suffered on account of credit risk or in terms of change in credit risk parameters. This paper focuses on impact of adverse macroeconomic development on probability of default of individuals.

In line with Vasicek (2002), the background for Basel II model of unexpected losses, we consider two types of credit risk factors affecting probability of default: systemic (global) and idiosyncratic (individual). Individual factors are specific for each transaction and are averaged out on diversified portfolio. Global factors are common to all transactions in the portfolio and can be interpreted as the macroeconomic conditions. Hence the macroeconomic variables are in the center of our interest and are viewed as drivers of key individual factors such as income and level of savings.

In literature, impact of the macroeconomic environment on credit risk has been studied from two perspectives. A bottom-up approach is based on data of individual clients while a top-down approach relies on aggregate data. This paper adopts the bottom-up approach similarly as Gross and Souleles (2002), who inspect stability of the credit risk models of the US households using the panel data set of credit cards. A top-down approach is more common in the literature as the aggregate data are more easily available: The time-series approach to modeling the arrears of UK individuals using the macroeconomic data is adopted by Whitley et al. (2004). A latent-factor model is developed by Jakubík (2007), which is then applied to the Czech banking sector.

When modelling credit stress in the Czech Republic the following issues are faced regarding macroeconomic variables:

- The transitive economy has not experienced regular economic cycle. The economic downturn in nineties of the last century was driven by specific economic conditions that are not likely to repeat.
- In last ten years the economic conditions gradually improve.
- Retail banking business has been dynamically developing within relatively short time from almost zero to its contemporary complexity.

Moreover, the portfolio is usually closely controlled by the Risk Management and hence we will develop a theoretical macroeconomic model rather than any statistical model. A clear story behind the results is also an advantage of this approach so that it can better serve "as a diagnostic tool to improve the institution's understanding of its risk profile". It is another important reason for stress testing as pointed out by Committee of European Banking Supervisors (2006).

The fundamental paper in the field of corporate default rate modelling and bond pricing is by Merton (1974). Our paper deals with modelling default rate of individual clients, which is not a topic frequently discussed in literature. We adopt the approach by Merton (1974) and modify it for the loans of individuals. Heřmánek et al. (2007) refer to "different sensitivities of corporations and households to the macroeconomic environment" which supports the need to build the model for individual clients differently.

Merton (1974) states that "on the maturity date, the firm must either pay the promised payment to the bondholders or else the current equity will be valueless." He concludes that the firm will default if the value of equity is lower than the promised payment. This conclusion is based on the assumption that "the firm cannot issue any new senior (or of equivalent rank) claims on the firm nor can it pay cash dividends or do share repurchase prior to the maturity date of the debt."

If we adopted this approach to the case of individuals, we would state that the client defaults if and only if the value of his property (incl. the net present value of his future income) is lower than the promised payment. However, assumption that the client cannot issue any claims on his property is not so realistic as it is in the case of firms. Hence, we will abstract of the client's property and focus solely on his income and savings.

The paper is organized as follows: Section 2.2 states the assumptions: Default is defined in Sec. 2.2.1, client's parameters are connected with the macroeconomic variables in Sec. 2.2.2, and the possible alternative of modelling the minimum consumption is presented in Sec. 2.2.3. Results are presented in Sec. 2.3: main implications for stress testing in Sec. 2.3.1, the parametric distribution of time to default under the substantially simplifying assumptions in Sec. 2.3.3. The two selected topics concerned with implementation of the model in the bank under Basel II IRB approach are discussed in Sec. 2.4 Section 2.5 concludes.

2.2 Theoretical framework

2.2.1 Definition of default

Denote D_t the random event that the client is in default in time t. Assume that the client cannot borrow additional money. Default is assumed to occur when the client's savings (incl. the liquid assets) and income do not cover all his expenditures, namely installments and consumption,

$$D_t \equiv [(1 + r^{\mathrm{s}}_{t-1})s_{t-1} + i_t - a_t - c_t + i_t\varepsilon_t < 0], \tag{2.1}$$

where s_t denotes client's savings, r^{s}_t the interest rate from savings, i_t the mean income, a_t the sum of installments of all loans, c_t the mean consumption, and ε_t is the random variable with zero mean describing shocks to both income and consumption. Multiplication of ε_t by i_t explicitly describes dependence of magnitude of the income and consumption shocks on the client's income.

The money the client does not pay on installments nor consumes turn into savings for the next period, hence the dynamics of savings

$$s_t = (1 + r^{\mathrm{s}}_{t-1})s_{t-1} + i_t - a_t - c_t + i_t\varepsilon_t \tag{2.2}$$

and the definition of default (2.1) can be rewritten as

$$D_t \equiv (s_t < 0). \tag{2.3}$$

Development of probability of default in time can be computed using this definition from the dynamics of savings, described by Eq. (2.2).

For convenience, we define $D_{1,t}$ the random event that the default occured in any of the time moments between 1 and t, i.e.

$$D_{1,t} \equiv (D_1 \vee D_2 \vee \ldots \vee D_t). \tag{2.4}$$

2.2.2 Stress testing

In this section, we let the client's income, installments, and consumption depend on the macroeconomic variables.

The mean consumption c_t, referred to in Eq. (2.1) and (2.2), is defined

$$c_t = \gamma_0 P_t/P_0 + \gamma_1(i_t - a_t) + \gamma_2(1 + r^{\mathrm{s}}_{t-1})s_{t-1}, \tag{2.5}$$

where $\gamma_1 \in [0, 1]$ and $\gamma_2 \in [0, 1]$ denote the marginal propensity to consumption from the disposable income, resp. savings. Parameter γ_0 specifies the minimum subsistence level in time 0, i.e. the necessary expenditures such as costs of food, basic accomodation etc. These expenditures are assumed to grow from γ_0 in time 0 at the same rate as the price level P_t does.

The mean income of the client is assumed to grow in line with the average income in the economy I_t,

$$i_t = i_0 I_t/I_0. \tag{2.6}$$

Income in the economy I_t should be built on the wage index W_t and account for the rate of unemployment U_t. For example, unemployed people can be assumed to receive social benefits of 6 000 CZK, then $I_t = (1-U_t)W_t+6000U_t$.

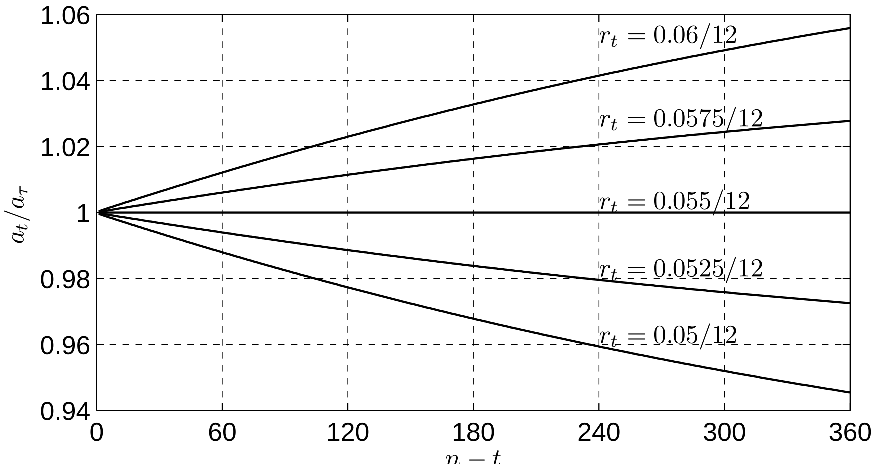

Figure 2.1: Results of equation (2.7) for $r_\tau = 0.055/12$ per month, i.e. 5.5% p.a.

For simplicity, we will assume exogenous I_t rather than separate W_t and U_t hereafter.

For consumer loans, annuity is constant over the whole life of the loan. For classic mortgage loans, annuity changes take place in agreed time of contract re-fix and depend on the interest rate as at the time of refix. Assume that re-fix takes place in time t. Maturity n, typically in months, will be expected to be fixed, i.e. only annuity, not time to maturity, changes in the time of re-fix. Denote τ time of previous interest rate fixation or loan granting. Then the residual outstanding amount in time t can be computed as the amount that remains to be paid with annuity a_τ and nominal interest rate r_τ for $n-t$ months. Then, new annuity is computed from this residual outstanding amount using the new loan interest rate r_t. The final formula is

$$a_t = a_\tau \cdot \frac{(1+r_\tau)^{n-t} - 1}{r_\tau(1+r_\tau)^{n-t}} \cdot \frac{r_t(1+r_t)^{n-t}}{(1+r_t)^{n-t} - 1}. \tag{2.7}$$

Magnitude of the change of loan annuity is depicted in Fig. 2.1 for various values of time to maturity and interest rate after re-fixation.

We will assume a constant additive gross margin of the loan interest rate r_t over the interest rate in the economy R_t. Hence if the loan is re-fixed in time t, its new interest rate is

$$r_t = r_0 - R_0 + R_t. \tag{2.8}$$

Remind that both r_t and R_t are assumed to be *per one period* rather than per annum!

2.2.3 Force of Habit: The Alternative Approach

In this section, we present a less traditional approach as an alternative to that discussed in Sec. 2.2.2. It leads to simpler results and may be more realistic for some economists.

We can use the per capita consumption C_t instead of the price index, i.e. set

$$P_t = C_t. \tag{2.9}$$

Justification of this approach is in the work of Abel (1990), who argues that the individual's utility depends on the ratio of his consumption to the lagged[1] cross-sectional average level of consumption. This approach is now popular in financial theory, see for example Campbell and Cochrane (1999).

To illustrate its reasonability on the practical example, consider a mortgage lasting 20 years. In the year 1988, which is exactly 20 years ago, people had no cell phones. In the year 2008, a cell phone belongs to the necessary equipment and people must buy it for their income. This rise of expenditures is not contained in the consumer price index. When cell phones (similarly as any other goods) entered CPI, only weights of the goods changed and it caused no immediate inflation. On the other hand, it was immediately reflected in the rise of per capita consumption.

Assume that on macroeconomic level, per capita consumption is the constant proportion of per capita income I_t,

$$C_t = \rho I_t. \tag{2.10}$$

Hypothesis of I_t/C_t constancy on the U.S. economy data spanning from Q1 1974 to Q1 1998 is discussed by Hayashi (2000), p. 648. The cointegration test of $\ln I_t$ and $\ln C_t$ (without the intercept term) is performed with the null hypothesis of no cointegration. His results reject the null hypothesis of no

[1]We will use the non-lagged average level of consumption because it results in a simpler formula but the lagged consumption could be employed as well.

cointegration on the subsample from Q1 1950 to Q4 1986, hence support cointegration. On the whole sample, the null hypothesis of no cointegration is not rejected. The reason is the declining personal saving rate from the mid 1980s. However, for our purposes, assumption of the constant personal saving rate is realistic enough.

2.3 Results

2.3.1 Stress testing

In this section, we provide implications of the assumptions stated in Sec. 2.2.2 and provide a Monte Carlo simulation of time-to-default distribution under different macroeconomic scenarios for a loan with and without re-fixation.

Equation (2.2) becomes after substitution from Eq. (2.5) and (2.6)

$$\begin{aligned}\frac{s_t}{i_t} &= (1-\gamma_1) + (1-\gamma_2)\frac{(1+r^{\mathrm{s}}_{t-1})s_{t-1}}{i_t} - (1-\gamma_1)\frac{a_t}{i_t} - \frac{\gamma_0}{i_t}\cdot\frac{P_t}{P_0} + \varepsilon_t \\ &= (1-\gamma_1) + (1-\gamma_2)(1+r^{\mathrm{s}}_{t-1})\frac{I_{t-1}}{I_t}\cdot\frac{s_{t-1}}{i_{t-1}} - (1-\gamma_1)\frac{a_t}{a_0}\cdot\frac{I_0}{I_t}\cdot\frac{a_0}{i_0} \\ &\qquad - \frac{I_0/P_0}{I_t/P_t}\cdot\frac{\gamma_0}{i_0} + \varepsilon_t. \end{aligned} \tag{2.11}$$

As $i_t > 0$, probability of default (PD) can be computed from this process of the savings-to-income ratio (SIR),

$$\mathrm{P}\left[D_t\right] = \mathrm{P}\left[s_t/i_t < 0\right]. \tag{2.12}$$

Hence, the factors rising SIR are also lowering PD and conversely.

Probability of default in time t given savings in time $t-1$ is due to Eq. (2.12) and (2.11)

$$\begin{aligned}\mathrm{P}\left[D_t\right] &= \mathrm{P}\left[s_t/i_t < 0\right] \\ &= \mathrm{P}\Big[\varepsilon_t < -(1-\gamma_1) - (1-\gamma_2)(1+r^{\mathrm{s}}_{t-1})\frac{I_{t-1}}{I_t}\cdot\frac{s_{t-1}}{i_{t-1}} \\ &\qquad + (1-\gamma_1)\frac{a_t}{a_0}\cdot\frac{I_0}{I_t}\cdot\frac{a_0}{i_0} + \frac{I_0/P_0}{I_t/P_t}\cdot\frac{\gamma_0}{i_0}\Big]. \end{aligned} \tag{2.13}$$

The most important client-specific factor contributing to the higher PD is the installment-to-income ratio (IIR) a_0/i_0, which is a measure of sensitivity to the stress of the installments and the nominal income. This factor can be managed by the bank in the granting process.

The other individual factor that increases PD is ratio of the minimum consumption to income γ_0/i_0, which determines sensitivity to the stress of real income I_t/P_t. This factor is less important relatively to IIR: For consumer loans, IIR determines sensitivity to the nominal income due to constancy of the installments. The nominal income is more volatile than the real income that affects the minimum consumption. For mortgages, the key role of IIR is sensitivity to the annuity stress, which is by far the most volatile risk driver.

The higher the savings-to-income ratio s_{t-1}/i_{t-1} the lower PD. However, it also measures sensitivity to the rise of income I_t/I_{t-1}, which *negatively* affects PD through this channel. The reason of this maybe controversial effect is the assumption of dependence of magnitude of income and supply shocks on the income. Hence, higher income means more volatile shocks. This factor is not very important because clients with high SIR have low PD and the income change is not as volatile as, e.g., the annuity for mortgage loans.

To sum up, impact of macroeconomic conditions is the following: Rise of interest rate R_t increases risk through the rise of a_t for re-fixed mortgages. Increased price level P_t rises the minimum consumption, hence also PD. Higher income in the economy I_t contributes to lower risk for the clients with low savings, who are the most riskiest ones. For the clients with very high savings, it rises PD due to the dependence of the magnitude of income shocks on the income. However, this group of clients is marginal.

In the rest of this section, we will provide examples of PD development. We will consider a client having a mortgage loan that is re-fixed each 12 months in the first example. In the second one, the same parameters will be used except for a constant annuity.

Trajectories for development of the economy are exogenous to our model. Several tentative scenarios are provided, covering beside the expected development the following unfavorable types of evolution: recession and growing interest rates. The scenarios are presented in Figure 2.2. They specify the annual interest rate $12R_t$, index of per capita income I_t (wages adjusted for the income of the unemployed) and price level index P_t. The time period is equal to one month.

Expected scenario is the point prediction of macroeconomic development. Recession will negatively affect incomes, i.e. wages decline and unemployment increases. Scenario with growth of interest rates is crucial for mortgage loans stress testing.

Monte Carlo simulations of development of the client's savings were performed via Eq. (2.11) using the independent Student's-t distributed disturbances with d degrees of freedom scaled by σ, i.e.

$$\varepsilon_t/\sigma \sim t(d). \tag{2.14}$$

This step was repeated 1 000 000 times and rates of default computed.

The following loan and client properties were used in the simulation:

$$\begin{aligned} i_0 &= 20\,000\,\text{CZK} & r_0 &= 0.055/12 \\ a_0 &= 7\,000\,\text{CZK} & r_t^{\text{s}} &= 0 \\ s_0 &= 5\,000\,\text{CZK} & \gamma_0 &= 5\,000\,\text{CZK} \\ n &= 360 & \gamma_1 &= 0.5 \\ \sigma &= 0.025 & \gamma_2 &= 0.5 \\ d &= 5 \end{aligned}$$

In the first example, loan interest rate is re-fixed each 12 months using Eq. (2.7). This is the case of the mortgage loans. Results can be seen in Figure 2.3. In the first graph there is the mortality function from the time of loan granting till the time t. The second graph describes probability of default in time t of the loans performing till time $t-1$. The third and fourth graph depicts development of annuity and loan interest rate in time.

In the second example, parameters of the loan are the same but annuity is constant over the whole life of the loan (resp. at least for the observed period). This may be the case of consumer loan or mortgage loan with a long fixation period. Note that maturity of the loan does not matter in this case. Default rates are depicted in Figure 2.4.

The increase of $\mathrm{P}\left[\mathrm{D}_t|\neg\mathrm{D}_{1,t-1}\right]$ after time $t = 0$ is caused by decreasing savings of part of the clients due to the adverse shocks. Contrary to this acts the effect of increasing average savings of those who did not default, which prevails in our example after the time of $t \approx 6$, when PD starts to decline.

While in the second example no serious impact of different scenarios on PD is observed, we have identified a serious risk of interest rate increase in the first example. It is caused by re-fixation of mortgage loans, when annuity substantially increases and causes the immediate growth of the probability of default of the non-defaulted loans in the next months.

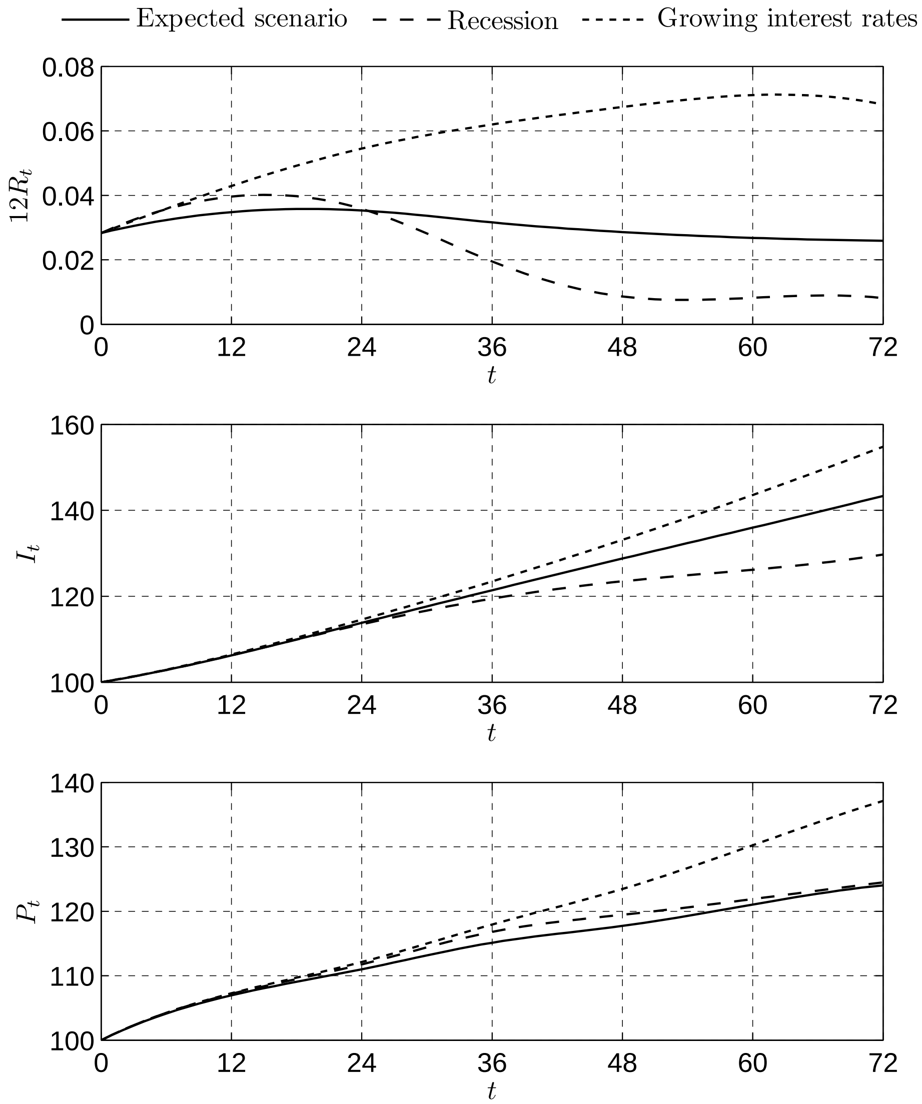

Figure 2.2: Scenarios of the macroeconomic development: interest rate p.a. $12R_t$, index of per capita income I_t and price level index P_t.

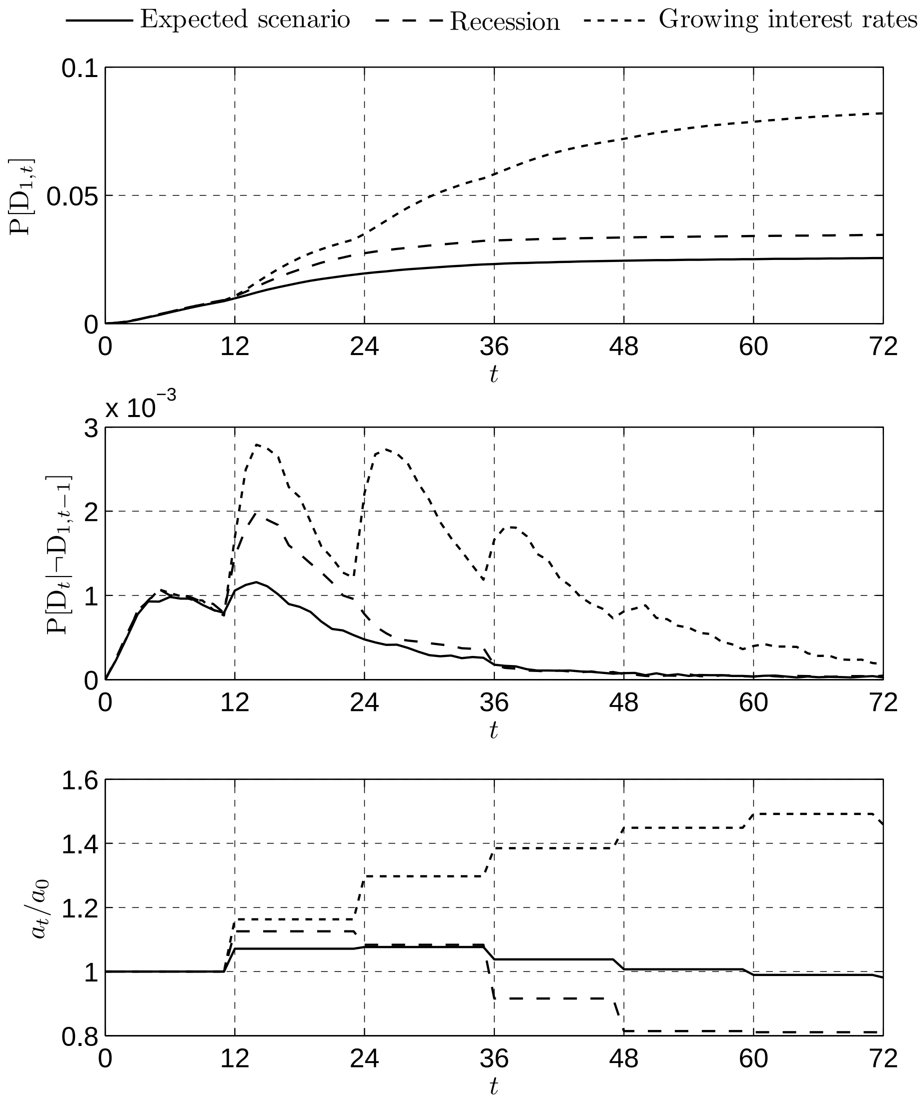

Figure 2.3: Annuities and PD's under different scenarios with re-fix each 12 months.

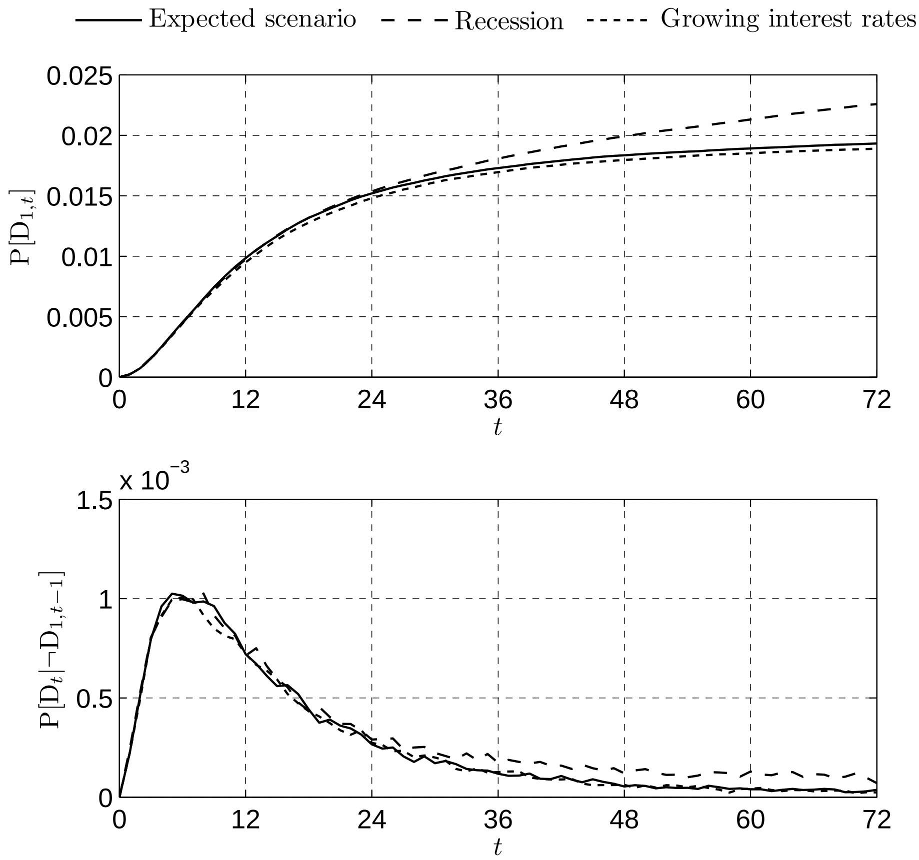

Figure 2.4: PD's under different scenarios with constant annuity.

2.3.2 The alternative approach

We briefly present implications of the assumptions from Sec. 2.2.3 in this section.

By equations (2.9) and (2.10), P_t is defined as

$$P_t = \rho I_t, \tag{2.15}$$

hence the price level is no longer necessary. Then, equation (2.11) becomes

$$\frac{s_t}{i_t} = (1-\gamma_1) + (1-\gamma_2)(1+r^{\mathrm{s}}_{t-1})\frac{I_{t-1}}{I_t}\cdot\frac{s_{t-1}}{i_{t-1}} - (1-\gamma_1)\frac{a_t}{a_0}\cdot\frac{I_0}{I_t}\cdot\frac{a_0}{i_0} - \frac{\gamma_0}{i_0} + \varepsilon_t. \tag{2.16}$$

and probability of default in time t given savings in time $t-1$ is

$$\mathrm{P}\left[D_t\right] = \mathrm{P}\Big[\varepsilon_t < -(1-\gamma_1) - (1-\gamma_2)(1+r^{\mathrm{s}}_{t-1})\frac{I_{t-1}}{I_t}\cdot\frac{s_{t-1}}{i_{t-1}} + (1-\gamma_1)\frac{a_t}{a_0}\cdot\frac{I_0}{I_t}\cdot\frac{a_0}{i_0} + \frac{\gamma_0}{i_0}\Big]. \tag{2.17}$$

In this approach, the minimum consumption γ_0 rises at the same rate as the income does, hence these two factors are compensated. The scenarios may include only interest rate and income (usually computed from wages and rate of unemployment).

2.3.3 Parametric distribution of time to default

In this section, we provide a by-product of our model: the explicit solution of the PDF of the time of first default under the considerably simplified asumptions and in continuous time. It is based on the assumptions of Sec. 2.2.1.

Consider a simplified case, when the mean income and the installments are independent of time, i.e. $i_t = i_0$ and $a_t = a_0$, and the interest rate from savings is $r^{\mathrm{s}}_t = 0$. The consumption is

$$c_t = \gamma_0 + \gamma_1(i_0 - a_0), \tag{2.18}$$

where γ_0 is the level of consumption under zero disposable income $i_0 - a_0$ and γ_1 is the marginal propensity to consumption.

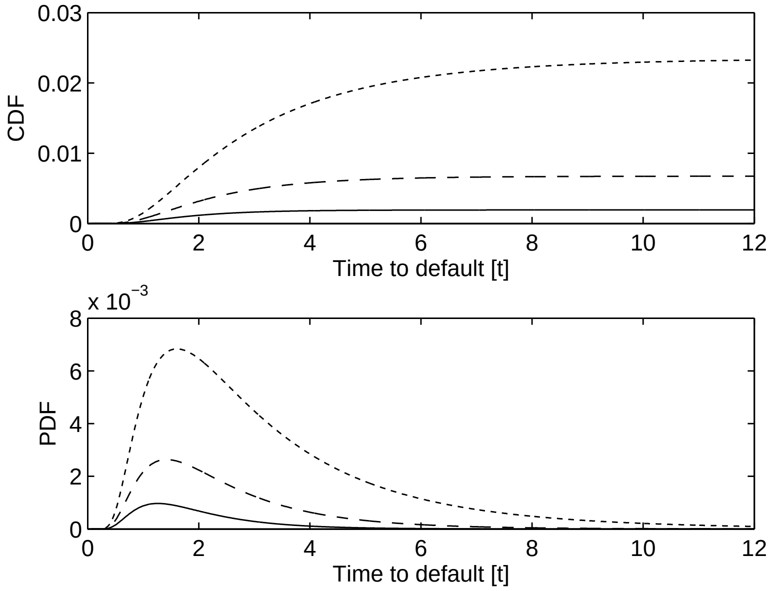

Figure 2.5: (a) CDF F of time to the first default from Eq. (2.20) and (b) the corresponding PDF for the common parameters $s_0 = 5\,000$, $i_0 = 20\,000$, $\gamma_0 = 5\,000$, $\gamma_1 = 0.5$, $\sigma = 0.1$, and the differing annuity $a_0 = 5\,000$ (solid line), $a_0 = 6\,000$ (dashed line), and $a_0 = 7\,000$ (dotted line).

After transition to the continuous time, Eq. (2.2) becomes after substitution of Eq. (2.18)

$$\mathrm{d}s_t = \nu\,\mathrm{d}t + i_0\sigma\,\mathrm{d}W_t, \tag{2.19}$$

where W_t is the Wiener process, $\sigma > 0$, and $\nu \equiv -\gamma_0 + (1-\gamma_1)(i_0 - a_0)$.

Then, according to Bielecki and Rutkowski (2002), p. 67, the time to the first default (i.e. the first passage time by s_t of the level 0) has an inverse Gaussian CDF

$$F(x) = \Phi(h_1(x)) + \exp\left\{-2\nu(i_0\sigma)^{-2}s_0\right\}\Phi(h_2(x)), \tag{2.20}$$

where Φ is CDF of the standard normal distribution, $s_0 > 0$, and

$$h_1(x) = \frac{-s_0 - \nu x}{i_0\sigma\sqrt{x}}, \qquad h_2(x) = \frac{-s_0 + \nu x}{i_0\sigma\sqrt{x}}. \tag{2.21}$$

Versions of CDF F and the corresponding PDF for various combinations of parameters are depicted in Fig. 2.5.

Note that the model in continuous time overestimates PD compared to the discrete model, simply because the condition of default is checked continuously instead of the discrete times only. This fact is most severe if the initial savings s_0 are close to 0: The time to default is then concentrated near $t = 0$.

Equation (2.20) can be used when modelling evolution of PD in time, e.g. for the purpose of computation of the standard risk costs. The curve should be fit to the observed default rates and the estimated parameters interpreted according to our model, rather than calibrate the parameters directly. For the practical purposes, default is defined as few installments past due, not only one due payment. The amount that is the boundary of default will be added to s_0, as marginal propensity to consumption from savings is zero in this section.

2.4 Implementation issues

In this section, we present some issues that are faced when implementing the presented model in the bank to fulfill requirements of Basel II IRB approach. If the reader is not interested in the practical implementation, he may skip this section.

Implementation of the model in a bank is a complicated issue, hence we discuss only two most frequent questions. However, the suggested solution does not necessarily have to be the only plausible one.

2.4.1 Definition of default

In Sec. 2.2.1, we assume that the client cannot borrow additional money and we define default using all the client's savings (incl. the liquid assets).

However, large number of clients can borrow additional money: They may draw their credit cards or authorized debits, be granted another loan from some financial institution, or borrow money from their family or friends. Savings of the client cannot often be determined precisely: They may have savings or other type of investments in a different institution.

This issues must be nevertheless solved by the PD model, which the bank must also maintain under Basel II IRB. Hence, it seems to be suitable to use the model proposed in this paper to simulate client's savings, credit turnover (income), debit turnover (consumption), and installments and use these variables as inputs to the PD model.

2.4.2 Transition from individual clients to the pools

The presented model is intended for stress testing of the individual client. It is computed by the Monte Carlo simulation which is time consuming, hence separate stress of all clients of the bank seems to be hardly possible.

However, the errors of estimation of PD of the individual clients are averaged out in the pool containing large number of clients. Hence, lower number of simulations for individual clients may be performed when interested in the stress of whole pools rather than the clients separately.

2.5 Conclusion

We proposed a dynamic macroeconomic model of individual client's PD. Scenarios of macroeconomic development and eventual fixation periods were specified exogenously.

The most important client-specific drivers of sensitivity to the stress were identified to be IIR and in case of mortgages also time to maturity. IIR determines sensitivity to the stress in the nominal income and annuity. Decrease of the nominal income generally increases PD of all portfolios of individuals. Annuity is the most inluential factor in the case of re-fixed mortgages: Its change depends on the interest rate stress and time to maturity. This conclusion was supported by the empirical simulation, where an important risk was found in the segment of mortgages under the scenario of growing interest rates.

This conclusion sends the clear message: Potential allowing for higher IIR and maturity length for mortgages in the granting process should be treated carefully.

An alternative approach based on the habit formation was presented. It does not require the exogenous price level and yields simpler results. Conclusions under this approach are similar.

As a by-product, explicit distribution of time to default in continuous time under the constant parameters was developed. It may be fitted to the default rate data in order to estimate the evolution of PD, when sufficiently long history is not available. This is needed, e.g., for the computation of the standard risk costs.

In the end, two questions regarding the practical implementation of the proposed model for stress testing under Basel II IRB approach were discussed.

Bibliography

A. Abel. Asset Prices under Habit Formation and Catching Up with the Joneses. *American Economic Review*, 80(2):38–42, 1990.

T. Bielecki and M. Rutkowski. *Credit Risk: Modeling, Valuation and Hedging.* Springer, 2002.

J. Campbell and J. Cochrane. By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior. *Journal of Political Economy*, 107(2):205–251, 1999.

Committee of European Banking Supervisors. Technical aspects of stress testing under the supervisory review process, 2006.

D. Gross and N. Souleles. An Empirical Analysis of Personal Bankruptcy and Delinquency, 2002.

F. Hayashi. *Econometrics.* Princeton University Press Princeton, 2000.

J. Heřmánek, M. Hlaváček, and P. Jakubík. Credit risk, credit growth models and stress testing. *CNB Economic Research Bulletin*, 1:25, 2007.

P. Jakubík. Macroeconomic Environment and Credit Risk. *Czech Journal of Economics and Finance*, 57(1-2):60–78, 2007.

R. Merton. On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journal of Finance*, 29(2):449–70, 1974.

O. Vasicek. The distribution of loan portfolio value. *Risk*, 15:160–162, 2002.

J. Whitley, R. Windram, and P. Cox. An empirical model of household arrears. *Bank of England Working Paper*, 214, 2004.

Essay 3

Correcting Predictive Models of Chaotic Systems

3.1 Introduction

Prediction of future development of many deterministic systems is often of practical importance. Success of this task depends on the quality of the model and accuracy of measurement of the initial conditions.

However, for chaotic systems, arbitrary deviation in the initial conditions leads to completely different trajectories. For such systems, the best long-term point prediction is the sample mean and the best predicted distribution is given by the natural measure of the system.

A suitable method, ensuring convergence of the point prediction to the sample mean in the long term, is the multi-step prediction. It finds a distinct functional relationship between the present and future state for each prediction horizon. This approach employing the local approximation is introduced by Farmer and Sidorowich (1987).

A better predictive model can be constructed when the theoretical properties of the system are reflected. However, such models are usually given by differential or difference equations, i.e. have the form of a one-step model. Also in the case of time-series models, the one-step model is more easily interpretable compared to the multi-step one, but unfortunately, point prediction performed by iterating a one-step model does not generally converge to the sample mean in the long term. The model loses correlation with the underlying system and its prediction error becomes higher than the error of the sample mean prediction.

Therefore it seems reasonable to specify some sort of transition between the one-step model and the natural measure (sample mean). In the literature, only a little attention has been given to this problem. Judd and Small (2000) take also a generally nonspecified (short-term) model and use the errors of its in-sample predictions to correct the out-of-sample predictions. However, they do not use the natural measure.

Our approach works as follows: For a particular prediction horizon, the distribution predicted by the short-term model is combined with the estimate of the natural measure. The uniquely given resulting distribution (resulting model) incorporates all information from the two sources. The Bayesian approach is employed but no additional information (e.g. a prior distribution) is required: The resulting model is derived from the source models under the reasonable assumptions. Note the difference that we model directly the distribution of the predicted variable, while in most applications [see e.g. Berliner (1991)], the Bayesian approach is used in the process of parameter estimation and predictions are made using the posterior distribution of parameters.

We show that the resulting model converges to the natural measure in the long term. We also prove that the point prediction by the resulting model dominates the point predictions by both source models in the sense of mean square error (MSE) for any horizon.

Since the distribution obtained by iterating the *nonlinear* one-step short-term model cannot usually be specified analytically, we propose a Monte Carlo method for its estimation. Similarly, real-world systems have usually non-typical natural measures. For such cases, we show the formula for the resulting model involving directly the past observations of the system instead of the analytically-specified natural measure.

Our approach can be employed in various fields of study, where iterated one-step predictions of chaotic systems are performed. Example of a large-scale system predicted in this paper is the U.S. unemployment rate. The small-scale dynamical systems are difficult to find in economics, hence we have adopted the physical one from Kantz and Schreiber (2004).

The paper is organized as follows: Section 3.2 states definitions of the chaotic system and natural measure (Sec. 3.2.1), assumptions about the underlying system (Sec. 3.2.2), the information available to our correcting procedure (Sec. 3.2.3), the properties of the resulting model (Sec. 3.2.4), and the model construction procedure (Sec. 3.2.5). Results are presented in Sec. 3.3: the unique distribution of the resulting model implied by the assumptions (Sec. 3.3.1), its long-term behavior (Sec. 3.3.2), and relationships among mean square errors of the models (Sec. 3.3.3). Section 3.4 focuses on practical implementation of the model: Monte Carlo simulation procedure of the short-term model (3.4.1) and formula using the nonparametrically specified natural measure (Sec. 3.4.2). In Sec. 3.5, examples are presented: Closed-form formula of the corrected AR(1) model's PDF under the Gaussian natural measure is derived (Sec. 3.5.1) and used for prediction of the U.S. unemployment rate (Sec. 3.5.2). The radial basis function network prediction of NMR laser data is performed by the Monte Carlo method and corrected using the nonparametric natural measure (Sec. 3.5.3).

3.2 Assumptions and definitions

3.2.1 General definitions

We state two general definitions used throughout the paper in this section.

Definition 1 (Natural measure) *The natural measure μ of the system given by transformation ϕ is defined as*

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} h(\phi^t(\boldsymbol{z}_0)) = \int_\Omega h(\boldsymbol{z})\, \mathrm{d}\mu(\boldsymbol{z}), \tag{3.1}$$

which is assumed to hold for every measurable function h and for the Lebesgue-almost every initial condition $\boldsymbol{z}_0$ form the state space Ω.

This is a definition similar to Eubank and Farmer (1997), p. 118, with a difference that we use a discrete time in the time average and just one natural measure is assumed to exist for a given system.

Under the term *chaotic*, we will understand *mixing* according to the following definition due to Lasota and Mackey (1994), p. 65.

Definition 2 (Mixing) *Let $(\Omega, \mathcal{A}, \mu)$ be a normalized measure space and $\phi : \Omega \to \Omega$ a measure-preserving transformation. ϕ is called* mixing *if*

$$\lim_{t\to\infty} \mu(A \cap \phi^{-t}(B)) = \mu(A)\mu(B) \qquad \textit{for all } A, B \in \mathcal{A}. \tag{3.2}$$

The transformation ϕ is measure-preserving with respect to the natural measure μ, i.e. $\mu(\phi^{-1}(A)) = \mu(A)$ for all $A, B \in \mathcal{A}$ and the given ϕ. The natural measure μ can thus be used in the definition of the mixing system.

The natural measure μ is a *probability measure* because $\mu(\Omega) = 1$. We can use the probabilistic notation such as $\mathrm{P}[\boldsymbol{z} \in A] \equiv \mu(A)$ and reformulate definition of the measure-preserving transformation ϕ as

$$\mathrm{P}[\phi(\boldsymbol{z}) \in A] = \mathrm{P}[\boldsymbol{z} \in A] \tag{3.3}$$

and definition of the mixing system as

$$\lim_{t\to\infty} \mathrm{P}\left[\boldsymbol{z} \in A \wedge \phi^t(\boldsymbol{z}) \in B\right] = \mathrm{P}[\boldsymbol{z} \in A]\,\mathrm{P}[\boldsymbol{z} \in B]\,. \tag{3.4}$$

Equation (3.3) implies $\mathrm{P}[\phi^t(\boldsymbol{z}) \in A] = \mathrm{P}[\boldsymbol{z} \in A]$ for all $t \in \mathbb{N}$, hence we can rewrite definition of the mixing system [Eq. (3.4)] as

$$\lim_{t\to\infty} \mathrm{P}\left[\boldsymbol{z} \in A \wedge \phi^t(\boldsymbol{z}) \in B\right] = \lim_{t\to\infty} \mathrm{P}[\boldsymbol{z} \in A]\,\mathrm{P}\left[\phi^t(\boldsymbol{z}) \in B\right]. \tag{3.5}$$

This equality says that the state of the system in time $t \to \infty$ does not depend on the set A of initial states and is distributed according to the natural measure.

3.2.2 Underlying system

In this section, we state assumptions about the underlying system, what is observed and what is predicted.

Assumption 1 (Underlying system) *Consider a mixing underlying system with the state vector $\boldsymbol{z}_t$ from the state space Ω and existence of its natural measure μ. Denote $\boldsymbol{y}_t$ the observed part of the state vector, the predicted vector $\boldsymbol{x}_t$ is part of $\boldsymbol{y}_t$. The subspace S of Ω contains all values of $\boldsymbol{x}_t$. The present time value is $t = 0$ and the observations are sampled at discrete times.*

We suppose that Assumption 1 holds throughout the article.

3.2.3 Available information

In this section, we state assumptions about the stochastic short-term model and the estimate of the natural measure, which are the only two sources of information for the correction procedure.

Consider that the short-term predictions are computed in discrete time moments $t \in \mathbb{N}$ with the discrete one-step model

$$\boldsymbol{y}_t = f(\boldsymbol{y}_{t-1}) + \boldsymbol{\eta}_t, \tag{3.6}$$

where function f is generally nonlinear, and the error term $\boldsymbol{\eta}_t$ has a zero mean. Function f is assumed to have a vector of estimated parameters denoted $\boldsymbol{\beta}_s$.

Estimation of parameters $\boldsymbol{\beta}_s$ of the short-term model is typically performed by minimizing the one-step prediction error on the historical data. For the large-scale systems, the short-term model can be seen as a local approximation of the underlying system.

Function f of the short-term model from Eq. (3.6) can be a theory-based or a time-series model, linear or nonlinear. Prominent example of the linear time-series method is the autoregressive moving average (ARMA) model. Among the nonlinear time-series models belong radial basis function networks (Kantz and Schreiber (2004), p. 212), neural networks (Kantz and Schreiber (2004), p. 213, Oliveira et al. (2000)), wavelet networks (Cao et al. (1995)), local linear models (Fan and Yao (2003), p. 20, Kantz and Schreiber (2004), p. 209) and many others. Review of the major estimation techniques with their one-step prediction performance evaluation can be found in Small et al. (2002).

Using the dynamic short-term model given by Eq. (3.6), we can compute the distribution of the short-term prediction of $\boldsymbol{x}_t$ for any horizon $t \in \mathbb{N}$. Note that p denotes a general PDF.

Assumption 2 (Short-term model) *Prediction of $\boldsymbol{x}_t$ by the short-term model has PDF $s_t(\boldsymbol{x}) \equiv p(\boldsymbol{x}|\boldsymbol{\beta}_s, t)$, where $t \in \mathbb{N}$ is an arbitrary prediction horizon and the vector $\boldsymbol{\beta}_s$ of estimated parameters is known. Assume there exists a function $s_\infty(\boldsymbol{x})$ and a function of time $K_t > 0$ such that $\lim_{t\to\infty} K_t s_t(\boldsymbol{x}) = s_\infty(\boldsymbol{x}) > 0$ for a.e. $\boldsymbol{x} \in S$.*

We will see later on that we will have to divide by s_∞ in the resulting formula. The introduction of K_t ensures that s_∞ is nonzero. If the limiting distribution of the short-term model is proper and supported on the whole S, function K_t may be chosen to be equal to 1 or arbitrary other positive constant. However, the limiting distribution is improper in many cases (e.g. for a Gaussian short-term model with variance growing to infinity, s_t converges to zero a.e.).

Denote P a general CDF.

Assumption 3 (Natural measure) *Assume we know an estimate of restriction of the natural probability measure μ to the subspace S of values of $\boldsymbol{x}_t$. It is represented by CDF M, vector of its estimated parameters is denoted $\boldsymbol{\beta}_m$, hence $M(\boldsymbol{x}) \equiv P(\boldsymbol{x}|\boldsymbol{\beta}_m)$.*

For the sake of simplicity, we will understand M rather than μ under the term 'natural measure' hereafter, unless explicitly stated differently. Contrary to the short-term model, the natural measure is independent of t. It represents only the limiting distribution, but does not describe any dynamics.

3.2.4 Resulting model

Our goal is to build a resulting model containing information from both the estimated parameters $\boldsymbol{\beta}_s$ of the short-term model and $\boldsymbol{\beta}_m$ of the natural measure:

Definition 3 (Resulting model) *Prediction of $\boldsymbol{x}_t$ by the* resulting model *is given by the CDF $R_t(\boldsymbol{x}) \equiv P(\boldsymbol{x}|\boldsymbol{\beta}_s, \boldsymbol{\beta}_m, t)$, where $t \in \mathbb{N}$ is an arbitrary prediction horizon.*

Assuming that there is no other information available, we want to build the resulting distribution only from the knowledge of values of $\boldsymbol{\beta}_s$ and $\boldsymbol{\beta}_m$. All eventual additional information should be incorporated into the short-term model or the natural measure.

3.2.5 Model construction procedure

The derivation of the resulting model is based on multiple usage of the Bayes theorem. In this section, we state three assumptions which enable us to compute the resulting model in the unique way.

In case we have no information about a present state of the underlying system, natural measure is the best prediction of its future state in an arbitrary time t. Hence, we use it as a prior:

Assumption 4 (Prior distribution) *The prior distribution of* $\boldsymbol{x}_t$ *is given by the natural measure, i.e.* $P(\boldsymbol{x}|\boldsymbol{\beta}_m, t) = P(\boldsymbol{x}|\boldsymbol{\beta}_m)$*, in the case we know the value of* $\boldsymbol{\beta}_m$*.*

The next assumption is:

Assumption 5 (Model independence) *Knowledge of estimated parameters* $\boldsymbol{\beta}_m$ *does not affect the distribution of* $\boldsymbol{\beta}_s$ *(conditional on* $\boldsymbol{x}$ *in time* t*),*

$$p(\boldsymbol{\beta}_s|\boldsymbol{\beta}_m, \boldsymbol{x}, t) = p(\boldsymbol{\beta}_s|\boldsymbol{x}, t). \tag{3.7}$$

Subjective joint distribution of parameter estimates $\boldsymbol{\beta}_s$ and $\boldsymbol{\beta}_m$ is given by the subjective joint distribution of the factors, which determine these estimates:

- historical data,
- extra information relevant for $\boldsymbol{\beta}_s$,
- extra information relevant for $\boldsymbol{\beta}_m$.

Assuming independence of these factors, the more important role of extra information, the lower dependence of $\boldsymbol{\beta}_s$ and $\boldsymbol{\beta}_m$. Note that conditioning by $\boldsymbol{x}$ restricts the subjective distribution of historical data to those trajectories believed to be crossing $\boldsymbol{x}$ in time t.

However, even in the case of no extra information, the dependence of $\boldsymbol{\beta}_s$ and $\boldsymbol{\beta}_m$ does not necessarily have to be strong:

- For the *large-scale* system (e.g. the world economy), the short-term model is a local approximation of the system, hence $\boldsymbol{\beta}_s$ depends on the part of the historical trajectory only, while $\boldsymbol{\beta}_m$ depends on the whole trajectory.
- For the *small-scale* system predicted using the complicated short-term model,[1] uncertainty about dynamics of the historical data (not affecting $\boldsymbol{\beta}_m$ but present in $\boldsymbol{\beta}_s$) is typically substantially higher than uncertainty about frequency of the observed states (affecting both $\boldsymbol{\beta}_m$ and $\boldsymbol{\beta}_s$).

[1] Under the term "complicated short-term model" we understand a highly nonlinear model or model with many parameters which can describe various types of dynamics, such as neural or radial basis function network.

Examples in Sec. 3.5.2 (large-scale system) and 3.5.3 (small-scale system) illustrate good results of the method for the case when $\boldsymbol{\beta}_s$ and $\boldsymbol{\beta}_m$ depend only on the historical data. Remind that final judgement about *plausibility of Assumption 5 depends solely on the user of the method because it deals with subjective distributions.*

Due to the assumption of the mixing underlying system, the states $\boldsymbol{z}_t \equiv \phi^t(\boldsymbol{z}_0)$ and $\boldsymbol{z}_0$ are independent for $t \to \infty$, see Eq. (3.5). Hence, vectors $\boldsymbol{x}_t$ and $\boldsymbol{y}_0$ are independent for $t \to \infty$ too because they are parts of $\boldsymbol{z}_t$ and $\boldsymbol{z}_0$. Analogically, independence of $\boldsymbol{x}_t$ and any of the historical states $\boldsymbol{y}_{-1}, \boldsymbol{y}_{-2}, \ldots$ for $t \to \infty$ can be shown. Hence, $\boldsymbol{\beta}_s$ might depend on $\boldsymbol{x}_t$ for $t \to \infty$ only due to some extra information but not due to its dependence on the historical observations. We assume that this is not our case:

Assumption 6 (Distant observation) *Knowledge of $\boldsymbol{x}_t$ for $t \to \infty$ does not affect the distribution of $\boldsymbol{\beta}_s$,*

$$\lim_{t\to\infty} p(\boldsymbol{\beta}_s|\boldsymbol{x}, t) = p(\boldsymbol{\beta}_s). \tag{3.8}$$

This assumption ensures convergence of R_t to M for $t \to \infty$ as discussed in Sec. 3.3.2.

3.3 Results

3.3.1 Resulting model formula

In this section, we derive CDF R_t of the resulting model using the known functions s_t, s_∞, and M.

We can observe both $\boldsymbol{\beta}_s$ and $\boldsymbol{\beta}_m$. The goal is to express distribution of $\boldsymbol{x}$ in a given future time t conditional on both of these parameter estimates. Let us take a prior distribution containing information about $\boldsymbol{\beta}_m$ and compute the posterior distribution (resulting model) using observation of $\boldsymbol{\beta}_s$:

$$\begin{aligned} \mathrm{d}R_t(\boldsymbol{x}) &= \mathrm{d}P(\boldsymbol{x}|\boldsymbol{\beta}_s, \boldsymbol{\beta}_m, t) \\ &= \frac{p(\boldsymbol{\beta}_s|\boldsymbol{\beta}_m, \boldsymbol{x}, t)\, \mathrm{d}P(\boldsymbol{x}|\boldsymbol{\beta}_m, t)}{\int_S p(\boldsymbol{\beta}_s|\boldsymbol{\beta}_m, \boldsymbol{\xi}, t)\, \mathrm{d}P(\boldsymbol{\xi}|\boldsymbol{\beta}_m, t)}. \end{aligned} \tag{3.9}$$

Unfortunately, the PDF and CDF involved in the RHS of the formula are unknown. In order to proceed, we employ Assumptions 3, 4, and 5 and get

$$\mathrm{d}R_t(\boldsymbol{x}) = \frac{p(\boldsymbol{\beta}_s|\boldsymbol{x}, t)\, \mathrm{d}M(\boldsymbol{x})}{\int_S p(\boldsymbol{\beta}_s|\boldsymbol{\xi}, t)\, \mathrm{d}M(\boldsymbol{\xi})}. \tag{3.10}$$

The term $p(\boldsymbol{\beta}_s|\boldsymbol{x},t)$ will be derived from s_t and s_∞ below by similar Bayesian reasoning as already used. We take a noninformative prior $p(\boldsymbol{x}|t)$ and compute the posterior distribution (short-term model) using observation of $\boldsymbol{\beta}_s$:

$$s_t(\boldsymbol{x}) \equiv p(\boldsymbol{x}|\boldsymbol{\beta}_s,t) = \frac{p(\boldsymbol{\beta}_s|\boldsymbol{x},t)p(\boldsymbol{x}|t)}{\int_S p(\boldsymbol{\beta}_s|\boldsymbol{\xi},t)p(\boldsymbol{\xi}|t)\,\mathrm{d}\boldsymbol{\xi}}. \tag{3.11}$$

As $p(\boldsymbol{x}|t) = \int_S p(\boldsymbol{x}|\boldsymbol{\beta}_m,t)p(\boldsymbol{\beta}_m|t)\,\mathrm{d}\boldsymbol{\beta}_m = \int_S p(\boldsymbol{x}|\boldsymbol{\beta}_m)p(\boldsymbol{\beta}_m)\,\mathrm{d}\boldsymbol{\beta}_m = p(\boldsymbol{x})$ due to Assumption 4 and the fact that the same parameter estimates $\boldsymbol{\beta}_m$ are used for any prediction horizon t, Eq. (3.11) becomes

$$s_t(\boldsymbol{x}) = \frac{p(\boldsymbol{\beta}_s|\boldsymbol{x},t)p(\boldsymbol{x})}{\int_S p(\boldsymbol{\beta}_s|\boldsymbol{\xi},t)p(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}}. \tag{3.12}$$

By computing limit of this equation, we obtain the relationship between $s_\infty(\boldsymbol{x})$ and $p(\boldsymbol{x})$,

$$s_\infty(\boldsymbol{x}) = \lim_{t\to\infty} K_t s_t(\boldsymbol{x}) = K_\infty p(\boldsymbol{x}), \tag{3.13}$$

where $K_\infty = \lim_{t\to\infty} K_t p(\boldsymbol{\beta}_s|\boldsymbol{x},t)/\int_S p(\boldsymbol{\beta}_s|\boldsymbol{\xi},t)p(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}$. We should show that K_∞ is really independent of $\boldsymbol{x}$: The only term in the K_∞ expression dependent on $\boldsymbol{x}$ in a given time t is $p(\boldsymbol{\beta}_s|\boldsymbol{x},t)$, which is however independent of x for $t \to \infty$ by Assumption 6.

Equation (3.13) thus states that the *noninformative prior is equivalent to the limiting distribution of the short-term model.*

The noninformative prior $p(\boldsymbol{x})$ may be improper, e.g. flat, and its choice depends on the particular application. For a list of noninformative priors see Yang and Berger (1996).

Now, only a few algebraic manipulations remain to be done to compute the distribution of the resulting model. Substitution of Eq. (3.13) into (3.12) results in

$$s_t(\boldsymbol{x}) = \frac{p(\boldsymbol{\beta}_s|\boldsymbol{x},t)s_\infty(\boldsymbol{x})}{\int_S p(\boldsymbol{\beta}_s|\boldsymbol{\xi},t)s_\infty(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}}. \tag{3.14}$$

After expressing $p(\boldsymbol{\beta}_s|\boldsymbol{x},t)$ from this equation and substitution into Eq. (3.10), the resulting model becomes

$$\mathrm{d}R_t(\boldsymbol{x}) = \frac{s_t(\boldsymbol{x})\mathrm{d}M(\boldsymbol{x})}{s_\infty(\boldsymbol{x})} \frac{\int_S p(\boldsymbol{\beta}_s|\boldsymbol{\xi},t)s_\infty(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}}{\int_S p(\boldsymbol{\beta}_s|\boldsymbol{\xi},t)\,\mathrm{d}M(\boldsymbol{\xi})}. \tag{3.15}$$

The above formula of the resulting model is almost final, it uses the known functions s_t, s_∞, and the natural measure M. The only remaining step is to eliminate the unknown multiplicative constant. We use the fact that $\int_S \mathrm{d}R_t = 1$ as R_t is a CDF. Equation (3.15) can now be simplified as follows

$$\mathrm{d}R_t(\boldsymbol{x}) = \frac{s_t(\boldsymbol{x})}{s_\infty(\boldsymbol{x})}\,\mathrm{d}M(\boldsymbol{x}) \left(\int_S \frac{s_t(\boldsymbol{\xi})}{s_\infty(\boldsymbol{\xi})}\,\mathrm{d}M(\boldsymbol{\xi})\right)^{-1}. \tag{3.16}$$

The sufficient condition for R_t to exist is that for a.e. $x \in S$: $s_t(\boldsymbol{x})/s_\infty(\boldsymbol{x})$ is bounded and positive.

Note that if there exists PDF m, such that $m(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \mathrm{d}M(\boldsymbol{x})$, then there exists also PDF of the resulting model

$$r_t(\boldsymbol{x}) = \frac{s_t(\boldsymbol{x})m(\boldsymbol{x})}{s_\infty(\boldsymbol{x})} \left(\int_S \frac{s_t(\boldsymbol{\xi})m(\boldsymbol{\xi})}{s_\infty(\boldsymbol{\xi})}\,\mathrm{d}\boldsymbol{\xi}\right)^{-1}. \tag{3.17}$$

3.3.2 Long-term predictions

In this section, we inspect the behavior of the resulting model for $t \to \infty$.

By Eq. (3.16), for any bounded continuous function g,

$$\lim_{t\to\infty} \int_S g(\boldsymbol{x})\,\mathrm{d}R_t(\boldsymbol{x}) = \lim_{t\to\infty} \int_S \frac{g(\boldsymbol{x})K_t s_t(\boldsymbol{x})}{s_\infty(\boldsymbol{x})}\,\mathrm{d}M(\boldsymbol{x}) \times \left(\int_S \frac{K_t s_t(\boldsymbol{\xi})}{s_\infty(\boldsymbol{\xi})}\,\mathrm{d}M(\boldsymbol{\xi})\right)^{-1}. \tag{3.18}$$

This equation can be simplified using the fact that $\lim_{t\to\infty} K_t s_t(\boldsymbol{x})/s_\infty(\boldsymbol{x}) = 1$ a.e., implied by Assumption 2. Under relatively mild additional assumptions, it can be proven by the Lebesgue dominated convergence theorem that $\lim_{t\to\infty} \int_S [K_t s_t(\boldsymbol{\xi})/s_\infty(\boldsymbol{\xi})]\,\mathrm{d}M(\boldsymbol{\xi}) = 1$ [see Billingsley (1986), p. 213]. Then,

$$\lim_{t\to\infty} \int_S g(\boldsymbol{x})\,\mathrm{d}R_t(\boldsymbol{x}) = \int_S g(\boldsymbol{x})\,\mathrm{d}M(\boldsymbol{x}), \tag{3.19}$$

hence the resulting model converges in distribution to the natural measure for $t \to \infty$.

3.3.3 Mean square error

In this section, we show that the resulting model dominates both the short-term model and natural measure in the sense of MSE for any prediction horizon.

Assume that all the available relevant information is contained in the parameter estimates $\boldsymbol{\beta}_s$ and $\boldsymbol{\beta}_m$. If we define the point prediction by the resulting model as

$$\hat{\boldsymbol{x}}_t \equiv \int_S \boldsymbol{x}\, \mathrm{d}R_t(\boldsymbol{x}) = \mathrm{E}\left[\boldsymbol{x}|\boldsymbol{\beta}_s, \boldsymbol{\beta}_m, t\right], \tag{3.20}$$

then its optimality in the MSE sense follows from Papoulis (1991), p. 175.

3.4 Computational issues

3.4.1 Monte Carlo predictions

We propose a method for computing the PDF of the short-term model by Monte Carlo simulations. This is particularly useful for the nonlinear short-term models, where the closed-form formula of the distribution is often difficult to obtain.

Consider a discrete-time one-step model described by Eq. (3.6),

$$\boldsymbol{y}_t = f(\boldsymbol{y}_{t-1}) + \boldsymbol{\eta}_t. \tag{3.21}$$

We can see that given the state vector $\boldsymbol{y}_{t-1}$, there is no problem with determination of the PDF of $\boldsymbol{y}_t$,

$$p(\boldsymbol{y}_t|\boldsymbol{y}_{t-1}) = \psi_t\big(\boldsymbol{y}_t - f(\boldsymbol{y}_{t-1})\big), \tag{3.22}$$

where ψ_t is PDF of $\boldsymbol{\eta}_t$.

However, given $\boldsymbol{y}_{t-2}$, the distribution of $\boldsymbol{y}_t$ is not so easy to compute anymore because in the formula

$$\boldsymbol{y}_t = f\big(f(\boldsymbol{y}_{t-2}) + \boldsymbol{\eta}_{t-1}\big) + \boldsymbol{\eta}_t, \tag{3.23}$$

the random variable $f(\boldsymbol{y}_{t-2}) + \boldsymbol{\eta}_{t-1}$ is the argument of the nonlinear function f. Prediction more than two steps ahead is even more difficult.

When we do not insist on the analytical formula, PDF of $\boldsymbol{y}_t$ given $\boldsymbol{y}_{t-j}$, $j \in \mathbb{N}$, can be generally expressed as

$$p(\boldsymbol{y}_t|\boldsymbol{y}_{t-j}) = \mathop{\mathrm{E}}_{\boldsymbol{y}_{t-1}}\left[\psi_t\big(\boldsymbol{y}_t - f(\boldsymbol{y}_{t-1})\big)|\boldsymbol{y}_{t-j}\right], \tag{3.24}$$

where $\mathrm{E}_{\boldsymbol{y}}$ denotes the mean value with respect to $\boldsymbol{y}$, i.e. $\mathrm{E}_{\boldsymbol{y}}\left[\psi(\boldsymbol{y})\right] \equiv \int_s \psi(\boldsymbol{y})\, \mathrm{d}\boldsymbol{y}$. This equation is a generalization of Eq. (3.22).

Our goal is to compute distribution of the short-term model s_t, i.e. the distribution of $\boldsymbol{x}_t$ given $\boldsymbol{y}_0$. As $\boldsymbol{x}_t$ is simply part of $\boldsymbol{y}_t$ (by Assumption 1), we can easily define[2] the function f_x and the random vector $\boldsymbol{\varepsilon}_t$ using f and $\boldsymbol{\eta}_t$, such that $\boldsymbol{x}_t = f_x(\boldsymbol{y}_{t-1}) + \boldsymbol{\varepsilon}_t$. Then

$$s_t(\boldsymbol{x}) = \operatorname*{E}_{\boldsymbol{x}_t^*}\left[\theta_t\left(\boldsymbol{x} - \boldsymbol{x}_t^*\right)|\boldsymbol{y}_0\right], \tag{3.25}$$

where $\boldsymbol{x}_t^* \equiv f_x(\boldsymbol{y}_{t-1})$ and θ_t is PDF of $\boldsymbol{\varepsilon}_t$. The Monte Carlo approximation of s_t is straightforward,

$$\hat{s}_t(\boldsymbol{x}) = \frac{1}{N}\sum_{i=1}^{N}\theta_t\left(\boldsymbol{x} - \boldsymbol{x}_{i,t}^*\right), \tag{3.26}$$

where $\boldsymbol{x}_{i,t}^*$ is the i-th simulation of $\boldsymbol{x}_t^*$ given $\boldsymbol{y}_0$. When choosing the N, the objective is that the approximation must be sufficiently good in the region, where the natural measure is concentrated.

As the approximation of $\hat{s}_\infty$, we take $\hat{s}_T$ with sufficiently high T, i.e. $\hat{s}_\infty \equiv \hat{s}_T$. Since θ_t is always proper and N finite, $\hat{s}_\infty$ is always proper (even if s_∞ is improper).

We have described the algorithm for the general distribution of $\boldsymbol{\varepsilon}_t$. Now, assume the frequent case of a scalar x_t in the subspace $S = \mathbb{R}$ and a Gaussian error $\varepsilon_t \sim N(0, \sigma^2)$. It can be shown that $\hat{s}_t/\hat{s}_\infty$ is bounded in $\mathbb{R}$ iff $\min_j\{x_{j,T}^*\} \leq x_{i,t}^* \leq \max_j\{x_{j,T}^*\}$ for all i, which is a sufficient condition for existence of $\hat{R}_t$ (approximation of R_t). When ε_t has a scaled Student's t-distribution, $\hat{s}_t/\hat{s}_\infty$ is bounded, hence $\hat{R}_t$ is guaranteed to exist.

3.4.2 Nonparametric natural measure

The natural measure can often be nonstandard and thus difficult to express parametrically. This is the case especially of the multivariate distribution: We are often able to construct its marginals but the problem is in specifying the whole joint distribution.

In this section, we describe a possibility to use historical values of $\boldsymbol{x}_t$ directly, without the need of specifying the natural measure in a parametric form. However, we must keep in mind that the nonparametric approach is more data-demanding.

[2] Function f_x takes the result $\boldsymbol{y}_t$ returned by f and restricts it to its part $\boldsymbol{x}_t$. Similarly, ε_t is part of $\boldsymbol{\eta}_t$.

Let CDF M be defined analogically to Eq. (3.1) as an approximation of the restriction of μ to the state subspace S of Ω along the typical trajectory of observed historical states $\boldsymbol{x}_{-T+1}, \ldots, \boldsymbol{x}_0$,

$$\frac{1}{T}\sum_{t=-T+1}^{0} h(\boldsymbol{x}_t) = \int_S h(\boldsymbol{x})\,\mathrm{d}M(\boldsymbol{x}), \tag{3.27}$$

which is assumed to hold for every measurable function h.

Using Eq. (3.16), we get for any measurable function g,

$$\int_S g(\boldsymbol{x})\,\mathrm{d}R_t(\boldsymbol{x}) = \left(\int_S \frac{g(\boldsymbol{x})s_t(\boldsymbol{x})}{s_\infty(\boldsymbol{x})}\,\mathrm{d}M(\boldsymbol{x})\right)\left(\int_S \frac{s_t(\boldsymbol{\xi})}{s_\infty(\boldsymbol{\xi})}\,\mathrm{d}M(\boldsymbol{\xi})\right)^{-1}. \tag{3.28}$$

[E.g. for calculation of the mean of the resulting model, we choose $g(\boldsymbol{x}) = \boldsymbol{x}$.] Equation (3.28) can be expressed using Eq. (3.27) as

$$\int_S g(\boldsymbol{x})\,\mathrm{d}R_t(\boldsymbol{x}) = \left(\frac{1}{T}\sum_{i=-T+1}^{0} \frac{g(\boldsymbol{x}_i)s_t(\boldsymbol{x}_i)}{s_\infty(\boldsymbol{x}_i)}\right)\left(\frac{1}{T}\sum_{i=-T+1}^{0} \frac{s_t(\boldsymbol{x}_i)}{s_\infty(\boldsymbol{x}_i)}\right)^{-1}. \tag{3.29}$$

Hence, only the short-term model and past observations are needed in this case as the sources for the correction procedure.

3.5 Examples

3.5.1 AR(1) model

In this example, we demonstrate the way of computation of the resulting model on the situation, when both the distribution of the short-term model and the natural measure are given analytically. The resulting model is derived for general parameters but no particular underlying system is assumed. It will be applied on the U.S. unemployment rate in Sec. 3.5.2.

Let the natural measure be Gaussian $N(\mu_m, \sigma_m^2)$ and the short-term model be AR(1),

$$x_t = a x_{t-1} + \varepsilon_t \tag{3.30}$$

with $a \in (0, \infty)$, $t \in \mathbb{N}$, and $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$, i.i.d. The prediction of the short-term model in time t is distributed $N(\mu_{s,t}, \sigma_{s,t}^2)$, where

$$\mu_{s,t} = a^t x_0, \tag{3.31}$$

$$\sigma_{s,t}^2 = \begin{cases} \sigma_\varepsilon^2 \frac{1-a^{2t}}{1-a^2}, & a \in (0,\infty)\setminus\{1\}, \\ t\sigma_\varepsilon^2, & a = 1. \end{cases} \tag{3.32}$$

After specification of both sources for the correcting procedure, we compute s_∞ according to Assumption 2. We choose K_t in order to obtain a simple formula for s_∞ (but it could be specified in many different ways),

$$K_t = \begin{cases} 1, & a \in (0,1), \\ \sqrt{2\pi}\sigma_{s,t}, & a = 1, \\ \sqrt{2\pi}\sigma_{s,t} \exp\left[\mu_{s,t}^2/(2\sigma_{s,t}^2)\right], & a \in (1,\infty). \end{cases} \tag{3.33}$$

Using the known PDF s_t and the above-defined function of time K_t, we can compute

$$\begin{aligned} s_\infty(x) &= \lim_{t\to\infty} K_t s_t(x) \\ &= \begin{cases} \varphi(x/\sigma_{s,\infty})/\sigma_{s,\infty}, & a \in (0,1), \\ 1, & a \in [1,\infty), \end{cases} \end{aligned} \tag{3.34}$$

where φ is PDF of $N(0,1)$ and $\sigma_{s,\infty} = \sigma_\varepsilon^2/(1-a^2)$.

Let us substitute s_t, s_∞, and PDF m concerning to the natural measure into Eq. (3.17) to compute PDF r_t of the resulting model. After some algebra, it can be calculated that r_t is PDF of $N(\mu_{r,t}, \sigma_{r,t}^2)$, where

$$\sigma_{r,t}^2 = \begin{cases} \frac{\sigma_{s,t}^2\sigma_m^2}{\sigma_{s,t}^2+\sigma_m^2(1-\sigma_{s,t}^2/\sigma_{s,\infty}^2)}, & a \in (0,1), \\ \frac{\sigma_{s,t}^2\sigma_m^2}{\sigma_{s,t}^2+\sigma_m^2}, & a \in [1,\infty), \end{cases} \tag{3.35}$$

$$\mu_{r,t} = \sigma_{r,t}^2 \frac{\mu_m\sigma_{s,t}^2 + \mu_{s,t}\sigma_m^2}{\sigma_{s,t}^2\sigma_m^2}. \tag{3.36}$$

It is worth noting that $\mu_{r,t}(a)$ and $\sigma_{r,t}^2(a)$ are continuous in $a = 1$ for any $t \in (0,\infty)$, hence the resulting model behaves similarly for a from the neighborhood of 1. In Fig. 3.1, we can see the alternatives of the short-term and resulting model for $a \in \{0.9, 1, 1.1\}$. The resulting models are similar for all three values of a regardless the fact that the AR(1) short-term model is stationary for $a \in (0,1)$ and nonstationary for $a \in [1,\infty)$. Note that we have used a continuous prediction horizon t for drawing the figure.

3.5.2 U.S. unemployment rate

Performance of the correction method is verified on the seasonally-adjusted monthly time series of the U.S. unemployment rate from 1948/01 till 2008/09,

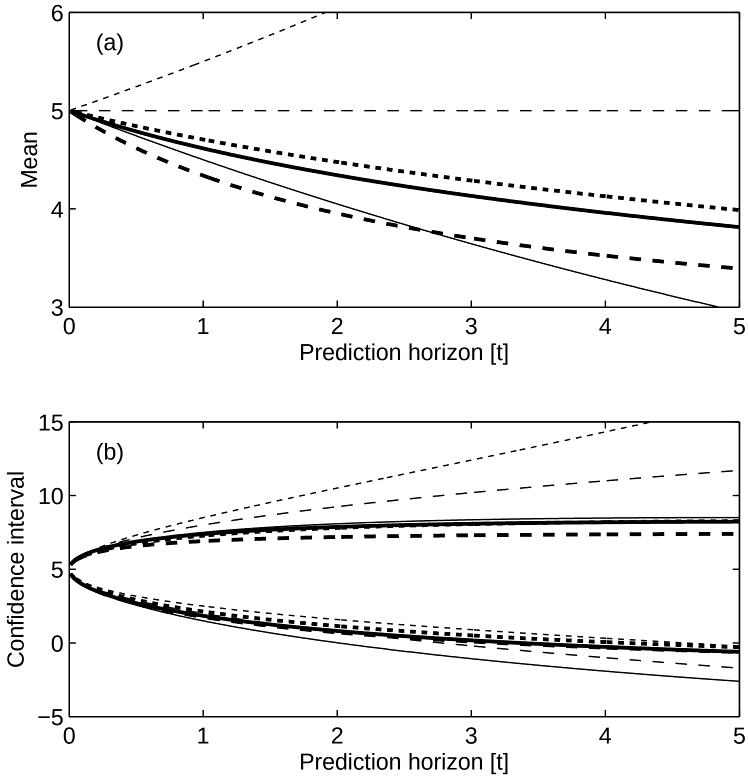

Figure 3.1: (a) Means and (b) two-sigma confidence intervals of the AR(1) short-term model (thin lines) and the resulting model (thick lines). Parameter a is 0.9 (solid lines), 1 (dashed lines), or 1.1 (dotted lines), the others are $x_0 = 5$, $\sigma_\varepsilon = 1.5$, $\mu_m = 2.5$, $\sigma_m = 2.5$.

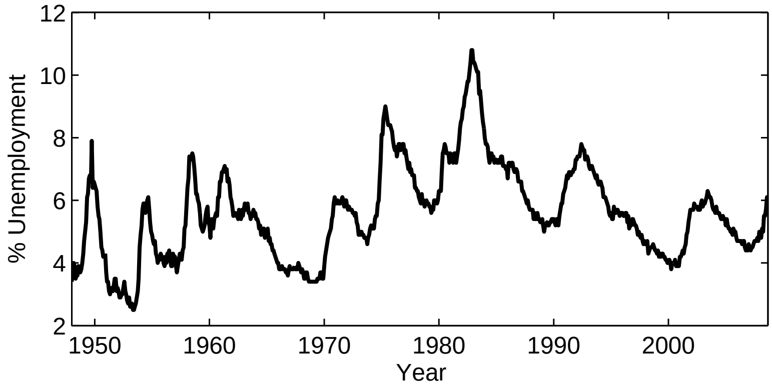

Figure 3.2: Time series of the U.S. unemployment rate used for demonstration of the correction method.

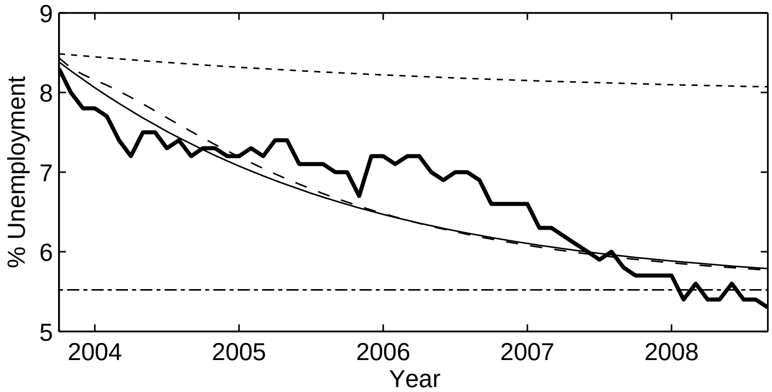

Figure 3.3: The U.S. unemployment rate (thick solid line) and its predictions using the short-term model (dotted line), the sample mean (dash-dot line), the resulting model from the natural measure given analytically (dashed line) and nonparametrically (solid line).

i.e. 729 observations (see Fig. 3.2). This is an example of a large-scale system.

The short-term model is the analytically-specified AR(1) model derived is Sec. 3.5.1. The natural measure is specified both parametrically assuming the Gaussian distribution and nonparametrically according to Sec. 3.4.2. RMSE's of the results are compared.

The computation is done using the open source R software, for reference see R Development Core Team (2008). The AR(1) model is fitted by function AR and predictions performed by function PREDICT.

For a parametric Gaussian natural measure, mean and variance are fitted to the data. Parameters of the resulting model are then computed using Eq. Eq. (3.35) and (3.36), adjusted for the generally nonzero mean of PDF s_∞.

Parameters of the resulting distribution, computed from the nonparametrical specification of the natural measure, are computed by Eq. (3.29).

Short-term model predictions are performed for the moving window, shifted by 1 observation from the very beginning to the end of the data set. The window contains 120 observations: The first 60 months are used as a sample for estimation of the model, the next 60 months are predicted. Hence, for each of the 610 positions of the moving window and a given prediction horizon $t = 1, \ldots, 60$, we obtain one realization of the prediction error for the computation of RMSE. The short-term model can be seen as a local approximation of the large-scale underlying system. On the other hand, all the data from the beginning of the time series till the end of the sample are used for estimation of the natural distribution.

Example of prediction is depicted in Fig. 3.3, RMSE of the predictions can be found in Fig. 3.4. We can see that although the short-term model dominates the sample mean, it can be further improved by the correction method. Performance of the resulting model computed using the parametrically and non-parametrically specified natural measure is very similar.

These results show the possibility to use the correction method for the large-scale systems and a short-term model estimated from the last observations. The data used for estimation of the natural measure must be assumed to be distributed similarly as the future observations.

3.5.3 NMR laser data

Due to the lack of small-scale systems with enough data in economics, we inspect the performance of the correction procedure on the time series of a signal measured in the NMR laser experiment (described in Flepp et al. (1991) and used for demonstration of nonlinear time-series methods in Kantz and Schreiber (2004)). The data set contains 39883 observations, which is enough to demonstrate the correction approach on the complicated nonlinear short-term model.

The short-term model is simulated by the Monte Carlo method, presented in Sec. 3.4.1, the natural measure is specified nonparametrically according to Sec. 3.4.2. For comparison, the deterministic version of the short-term model is also estimated.

The computation is performed using routines of the TISEAN package (`http://www.mpipks-dresden.mpg.de/~tisean`), an open source software implementing many nonlinear time series methods. For reference, see Hegger et al. (2000).

The short-term model is in the form of Eq. (3.21), where f is a radial basis function network (RBF). It is generally specified as $f(\boldsymbol{y}) = \sum_{i=1}^{n} w_i \Phi(\|\boldsymbol{y} - \boldsymbol{c}_i\|)$, where w_i and $\boldsymbol{c}_i$ denote weights and centers. The parameters of the RBF are set according to Kantz and Schreiber (2004), p. 246, Example 12.5: Lorentzian function $\Phi(r) = 1/(1 + r^2/a^2)$, $a = 1770$ and $n = 140$. The centers are distributed according to the algorithm of the RBF routine, which was modified to compute also the corrected RBF predictions. Errors are assumed to be independent and have a scaled Student's t-distribution with 3 degrees of freedom (more than 2 degrees imply a finite variance). The number of Monte Carlo iterations for both $\hat{s}_t$ and $\hat{s}_\infty$ is $N = 1000$, the infinite time horizon is set to $T = 200$.

As the short-term model in the form of Eq. (3.21) does not account for measurement errors, the data set is first cleared by the locally projective noise reduction, implemented in the PROJECT routine. The parameters are set according to Kantz and Schreiber (2004), p. 185, Example 10.3.

In order to compute the root mean square error (RMSE), predictions are performed for the moving window, shifted by 1 observation from the very beginning to the end of the data set. The window contains 4050 observations: The first 4000 of them are used as a sample, the next 50 steps are predicted. Hence, 35834 distinct positions of the moving window exist.

All the 4000 observations from the sample are used to estimate the deterministic version of the short-term model and to specify the natural measure according to Sec. 3.4.2. When estimating the short-term model, the sample is further split into the training set (3600 observations) and a test set (400 observations), used for estimation of variance of the error term.

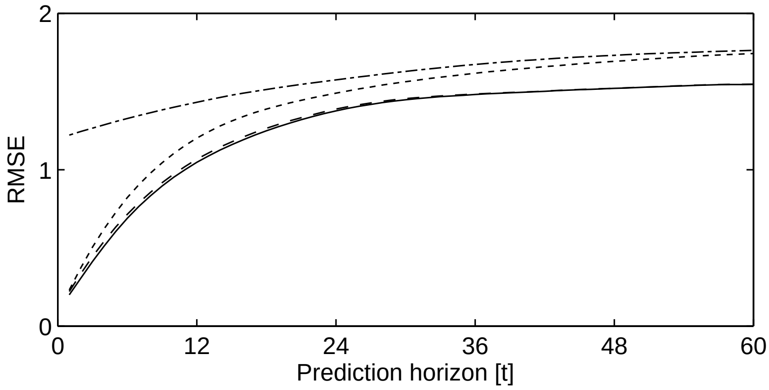

Figure 3.4: RMSE of the U.S. unemployment rate predictions using the short-term model (dotted line), the sample mean (dash-dot line), the resulting model from the natural measure given analytically (dashed line) and nonparametrically (solid line).

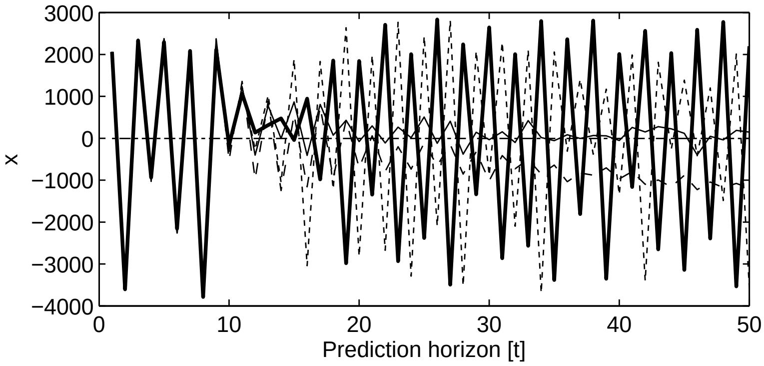

Figure 3.5: Example of the NMR laser data (thick solid line) and its predictions using the short-term model (dashed line), its deterministic version (dotted line), the sample mean (dash-dot line) and the resulting model (solid line).

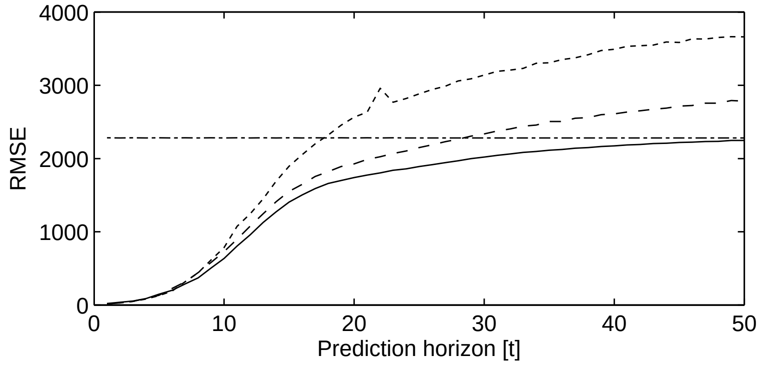

Figure 3.6: RMSE of the NMR laser data predictions using the short-term model (dashed line), its deterministic version (dotted line), the sample mean (dash-dot line) and the resulting model (solid line).

Example of the predictions can be seen in Fig. 3.5. Paths of both the short-term model and the resulting model are similar to the true data up to $t \approx 12$. Then the models lose correlation with the chaotic underlying system. This shortcoming is reduced by the resulting model, which converges to the sample mean, i.e. mean of the distribution specified by the natural measure. Note the decreasing variability of the (stochastic) short-term model compared to the unchanged dynamics of its deterministic version.

RMSE of the predictions is depicted in Fig. 3.6. (The spike on the RMSE curve of the deterministic version of the short-term model is caused by an outlier in one position of the moving window.) These results confirming the superiority of the resulting model to both the short-term model and the natural measure are in concordance with Sec. 3.3.3.

3.6 Conclusion

We have derived the unique formula for the resulting model combining the short-term model and natural measure.

We have proven two desirable properties: First, the resulting model converges to the natural measure in the long term. Second, the mean of the resulting model dominates the means of both the source models in the sense of MSE.

In real-world problems, it is often impossible to express both the PDF of the short-term model and the natural measure analytically. Therefore, we presented a Monte Carlo approach allowing to compute the distribution of the short-term model numerically and a nonparametric approach enabling to derive the resulting model directly from the short-term model and past observations.

Finally, three examples were presented. In the first one, the closed-form formula for the resulting model was derived analytically for an AR(1) short-term model and a Gaussian natural measure. This result was applied to the time series of U.S. unemployment rate and RMSE of the predictions computed. The resulting model was better for all prediction horizons than the AR(1) short-term model despite the fact that the short-term model dominated the sample mean. The third example demonstrated the correction procedure on the the NMR laser dataset. The short-term model was approximated by the Monte Carlo method, the natural measure was specified nonparametrically. Again, for all prediction horizons, the presented correction procedure turned out to be superior (in the sense of MSE) to both the short-term model and the sample mean.

Bibliography

L. Berliner. Likelihood and bayesian prediction of chaotic systems. *J. Am. Stat. Assoc.*, 86:938, 1991.

P. Billingsley. *Probability and Measure*. Wiley, New York, 1986.

L. Cao, Y. Hong, H. Fang, and G. He. Predicting chaotic time series with wavelet networks. *Phys. D*, 85:225, 1995.

S. Eubank and J. Farmer. Introduction to dynamical systems. In L. Lam, editor, *Introduction to Nonlinear Physics*, page 55, New York, 1997. Springer.

J. Fan and Q. Yao. *Nonlinear Time Series*. Springer, New York, 2003.

J. D. Farmer and J. J. Sidorowich. Predicting chaotic time series. *Phys. Rev. Lett.*, 59:845, 1987.

L. Flepp, R. Holzner, E. Brun, M. Finardi, and R. Badii. Model identification by periodic-orbit analysis for nmr laser chaos. *Phys. Rev. Lett.*, 67:2244, 1991.

R. Hegger, H. Kantz, and T. Schreiber. Practical implementation of nonlinear time series methods: The tisean package. *Chaos*, 9:413, 2000.

K. Judd and M. Small. Towards long-term prediction. *Phys. D*, 136:31, 2000.

H. Kantz and T. Schreiber. *Nonlinear Time Series Analysis*. Cambridge University Press, Cambridge, 2004.

A. Lasota and M. Mackey. *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics*. Springer, 1994.

K. Oliveira, A. Vannucci, and E. Silva. Using artificial neural networks to forecast chaotic time series. *Phys. A*, 284:393, 2000.

A. Papoulis. *Probability, random variables, and stochastic processes*. McGraw Hill, New York, 1991.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2008. URL `http://www.R-project.org`.

M. Small, K. Judd, and A. Mees. Modeling continuous processes from data. *Phys. Rev. E*, 65:046704, 2002.

R. Yang and J. Berger. A catalogue of noninformative priors. Technical report, Purdue University, 1996.